INFORMATION FOR CONTRIBUTORS

GENERAL

All manuscripts should be addressed to the Editor, *Current Science*, P. B. No. 8001, C. V. Raman Avenue, Bangalore 560 080. Submission of an article will be held to imply that it has not been previously published and is not under consideration for publication elsewhere; and further, that if accepted, it will not be published elsewhere. *Three copies of contributions of all categories* are required, with a letter of transmittal giving (i) names and complete addresses of the authors and (ii) title of the contribution and the category in which it is submitted (see below).

Current Science is a multidisciplinary journal and therefore research and review papers of general significance that are written clearly and well organized will be given preference. All papers will be first assessed by a Reviewing Editor. Papers found unsuitable in terms of the overall requirements of the journal will be returned to the authors. The others will be sent for detailed review. *Both solicited and unsolicited material will be reviewed.* Authors of these papers will be notified of acceptance, rejection, or need for revision of the paper. Returned papers cannot be resubmitted.

Illustrations and other material to be reproduced from other publications must be properly credited; it is the authors' responsibility to obtain permission for reproduction (copies of letters of permission should be sent).

CATEGORIES OF MANUSCRIPT

General articles (not exceeding 5000 words) discuss current trends in research in a field that will be of interest to readers outside the field; interdisciplinary topics: science policy and administration; or some aspect of the application of science and technology to human needs or the impact of science and technology on society/ecosystems/life. They should include a summary not exceeding 100 words, introductory paragraph(s), brief subheads at appropriate places to point to what follows, illustrations that will help a general reader, and references.

Review articles (not exceeding 5000 words) are expected to survey and discuss recent developments in a field. They should be well focused and organized, and avoid a general, 'textbook' style.

Research articles (not exceeding 4000 words) should report research results of fairly major significance. They should include an abstract not exceeding 100 words, introductory paragraph(s), and brief subheads.

Research communications (not exceeding 2000 words) should contain important findings that are novel and of fairly broad interest. They should include a brief abstract and an introductory paragraph. Text should not be broken up under subheads.

Correspondence includes letters, not exceeding 500 words, that are of general interest to scientists. All letters cannot be published.

Scientific correspondence contains technical comments, including those on articles or communications published in *Current Science* within the previous six months. Letters may be reviewed and edited.

Research news articles are intended to inform nonspecialists about recently published advances or important findings discussed at a meeting. Authors should also send a copy of the paper(s) on which the article is based. Meeting reports should avoid merely listing brief accounts of topics discussed, and must convey to readers the significance of an important advance.

Research accounts articles are intended to be personalized reviews of research from the authors' own laboratory, based on a body of published work. The articles must provide appropriate background to the area in a concise introduction, which should also serve to place the author's work in proper perspective. Articles will normally not exceed 8 to 10 printed pages.

Opinion articles present views on issues related to science and scientific activity. **Commentary** articles should contain expository notes on issues related to science and scientific activity.

Book reviews. Unsolicited reviews will also be considered. Reviews that merely 'list' brief descriptions of the contents cannot be published. Reviews should have 'context' and convey some information about the subject of the book.

Historical commentary and notes inform readers about interesting aspects of personalities or institutions of science or about watershed events in the history/development of science; most are expected to relate to India. Illustrations are welcome. Brief items will also be considered.

MANUSCRIPT PREPARATION

Manuscripts should be typed double-spaced on one side of white bond paper (21 × 28 cm). The pages should be numbered consecutively, starting with the title page and through the text, reference list, tables and figure legends. The **title** should be brief, specific and amenable to indexing. Not more than five **keywords** should be indicated separately; these should be chosen carefully and must not be phrases of several words. **Summary** and **abstract** should not have more than 100 words and should convey the main point of the paper, outline the results and conclusions, and explain the significance of the results.

Text. All papers should have a brief introduction. The text should be intelligible to readers in different disciplines and technical terms should be defined. Tables and figures should be referred to in numerical order. All **symbols** and **abbreviations** must be defined, and used only when absolutely necessary. Superscripts and subscripts and ambiguous characters should be clearly indicated. **Units of measure** should be metric or, preferably, SI. Methods should, as far as possible, be described briefly in appropriate table and figure legends.

Figures. In the case of line drawings one set of originals (without any lettering) is sufficient, with two copies containing lettering. In the case of photographs good prints are required with each copy of the manuscript; photocopies are not acceptable. Line drawings should be roughly twice the final printed size. The correct orientation should be indicated if not clear.

Photomicrographs and other photographs that require it must have a scale bar, which should be defined clearly in the legend. Primary data should be submitted as far as possible (e.g. actual photographs of electrophoretic gels rather than idealized diagrams).

References should be numbered in the order in which they appear, first through the text and then through table and figure legends. The following are examples of ways of writing references.

1. Mukundan, T. and Kishore, K., *Curr. Sci.*, 1991, **60**, 355–362.
2. Constantine. G., in *Biology of Bats* (ed. Wimsatt. W. A), Academic Press, New York. 1970. vol 1, pp. 319–322.

Acknowledgements should be brief. Footnotes are not allowed except to identify the corresponding author if not the first.

Cover photographs. Good photographs (colour or black and white) that pertain to a submitted paper will be considered for use on the cover. Good *prints* and a legend should be submitted with the manuscript. In the case of a colour picture, a transparency will be required for printing if accepted.

PROOFS AND PUBLICATION

Two sets of galley proofs are sent to the corresponding author. A reprint order form accompanies the proofs.

Founded: 1932

ISSN: 0011-3891

CURRENT SCIENCE

Volume 69 Number 1

10 July 1995

3 In this issue

CORRESPONDENCE

5 **Fancy and facts in science and art**
B. M. Prasanna

6 **Distortion of Indian degrees**
S. C. Dutta Roy

NEWS

7 **Amending the Patents Act – An appeal from the National Working Group on Patent Laws**

RESEARCH NEWS

9 **Binding and release of polypeptides during groEl-mediated folding: Its cellular significance?**
Amitabha Chaudhuri

COMMENTARY

12 **Importance of basic research in the context of global change**
R. Chidambaram

14 **Census of India's biodiversity: Tasks ahead**
T. N. Khoshoo

SCIENTIFIC CORRESPONDENCE

17 **Record of opal phytoliths and associated palynofossils from Indian crude oils**
C. M. Misra and R. K. Saxena

21 **Massif-type anorthosite near Ayialur, Dindigul, Anna District, Tamil Nadu, South India**
M. Arumugam and R. Senthil Kumar

21 **A technique for axenic culture of *Puccinia romagnoliana* – A potential bioherbicide for purple nutsedge**
J. S. Bedi, Ravinder Kaur and S. S. Sokhi

23 **Polyphenolic components of flowers of *Helichrysum bracteatum***
R. Sivakumar, A. G. Ramachandran Nair and K. P. Madhusudhanan

GENERAL ARTICLES

25 **Reviving experimental physics in the universities – The IUC initiative**
S. Ramaseshan

29 **Modern science and technology: The public perception**
V. Govindarajulu

39 **Impact of extreme events on coastal zones and small islands in the context of climate change**
S. M. Kulshrestha

RESEARCH ARTICLES

44 **New pressure-induced phase changes in $CdWO_4$ from Raman spectroscopic and optical microscopic studies**
A. Jayaraman, S. Y. Wang and S. K. Sharma

48 **Analysis of a conformation-specific epitope of the alpha subunit of human chorionic gonadotropin: Study using monoclonal antibody probes**
N. Venkatesh, G. Nagaraja and G. S. Murthy

RESEARCH COMMUNICATIONS

57 **On minimum metallic diffusivity in disordered conductors**
N. Kumar and A. M. Jayannavar

58 **An *ab initio* topographical investigation on the molecular electrostatic potential of some chemical mutagens**
Apurba K. Bhattacharjee, Savita S. Pundlik and Shridhar R. Gadre

62 **Record of 1.66-Ga-old marine oxido-reducing conditions from an unusual Sargipali Pb–Cu deposit (India): A preliminary appraisal through petrogenesis**
Rajesh K. Vishwakarma

67 **Characterization of humic acids isolated from soils under mulberry vegetation**
R. Kar, R. N. Dutta and S. K. Majumder

BOOK REVIEWS

69 **Contributions to Operator Theory and its Applications,** *reviewed by* Rajendra Bhatia

69 **Science in Society: Some Perspectives,** *reviewed by* R. Parthasarathy

71 **Annual Review of Cell Biology 1994,** *reviewed by* Veena K. Parnaik

72 **Environment and Adaptive Biology of Plants,** *reviewed by* R. J. Ranjit Daniels

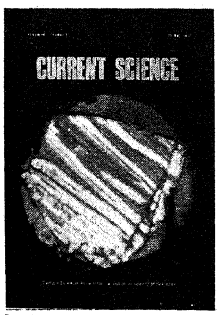
PERSONAL NEWS

73 **A pioneer in modern plasma physics,** *an obituary of Hannes Alfvén by* Abhijit Sen and P. K. Kaw

74 **Remembering Hannes Alfvén and his work**
Chanchal Uberoi

HISTORICAL COMMENTARY AND NOTES

76 **Professor S. N. Bose's contribution to quantum statistics – A critical review**
S. K. Das and S. Sengupta



COVER. Photomicrograph of a $CdWO_4$ crystal at high pressure. See page 44.

Indexed in CURRENT CONTENTS/GEOBASE/CHEMICAL ABSTRACTS

In this issue

Pressure-induced phase changes in $CdWO_4$

Studies in pressure-induced phase transformation in sheelite type ABO_4 compounds by Raman spectroscopy have demonstrated that such transformations are a rule in these systems. Do Wolframite type ABO_4 compounds exhibit pressure-induced transformation (PIT) is a logical question. A. Jayaraman *et al.* answer this question in their article 'New pressure-induced phase changes in $CdWO_4$ from Raman spectroscopic and optical microscopic studies' on **page 44**.

Employing a gasketted Mao–Bell type diamond cell, the behaviour of $CdWO_4$ (Wolframite structure) at high pressures have been studied up to 40 GPa with the help of high pressure Raman spectra. The results show that wolframite phase disappears above 30 GPa and reappears at 16 GPa on release of pressure. This transition is completely reversible at room temperature and the high pressure phase is not quenchable. It is a first order transition and is hysteritic.

Identifying the highest wave number peak with ν_1 mode of WO_6, the authors argue that the decrease of the frequency of this peak by about 120 cm^{-1} in the high pressure phase (as compared to the wolframite phase) is indicative of a truly octahedral W–O coordination in the high pressure phase. They point out that the description of octahedrally coordinated WO_6 as building blocks in wolframite $CdWO_4$ is a highly distorted picture. It should really be regarded as tetrahedral and that is also consistent with the Raman frequency of WO_4 in wolframite $CdWO_4$.

Comparing their results with the ν_1 mode of the octahedral WO_6 in Ba_2CaWO_6, they conclude that in the high pressure phase of $CdWO_4$, the coordination of W becomes truly octahedral. They suggest that the WO_6 octahedrons may form chains in the structure sharing four corners with the neighbouring octahedra or from zig-zag chains sharing the edges to preserve the W–O ratio. They point out the need for additional input from high pressure X-ray study to identify the other weak peaks of Raman spectrum.

Optical observations under microscopic (in unpolarized and crossed polarized light) with thick (~ 50 μm) as well as thin (~ 20 μm) samples of $CdWO_4$ revealed a remarkable PIT near 10 GPa. This is also shown in Raman spectra, in a less pronounced way, as small changes of slope in pressure vs Raman shift plot. This subtle transition in the wolframite phase is interpreted as the wolframite lattice 12/a taking on the C2/c symmetry of $HgWO_4$.

The authors show in the case of sheelite $CdMoO_4$, the sheelite to wolframite transition near 12 GPa. Near 25 GPa, $CdMoO_4$ exhibits another transition similar to that observed in $CdWO_4$.

Based on the above results, the authors suggest that the picture for the high pressure transition sequence in $A^{2+}B^{2+}O_4$ compounds is sheelite to wolframite or wolframite-like phase and then to a phase with truly octahedrally coordinated BO_6 units.

S. R. Rajagopalan



Enzymes, Pharmaceuticals & Industrial Chemicals Ltd., BOMBAY

EPIC is an Indo-German joint venture company listed on the Bombay Stock Exchange with a turnover of about Rs.40 crores. Its main manufacturing unit and R & D Laboratories are located at Lohop, about 70 kms. South of Bombay near Panvel and with a second unit at Taloja, near Bombay. The Company produces a broad line of speciality chemicals, bulk drugs, pesticides and chemical intermediates, some of them exclusively for its European collaborators.

Of the 150 employees of the Company over 100 are grade engineers and chemists from institutions like the Indian Institute of Technology, Bombay University, Tufts University and Harvard University.

The Company has several vacancies at its R & D Laboratories at Lohop.

RESEARCH SCIENTISTS

Organic Chemists having proven ability for independent research as evidenced by publications. Minimal qualification is M.Sc. with 10 years experience and consistently good academic record or Ph.D. with three years post doctoral experience in a University or Industry. The Scientists are expected to develop innovative and cost-effective processes and procedures for making various Organic Chemicals used as pesticides, drugs and intermediates. They should be capable and willing to work with their own hands in addition to guiding junior scientists and technicians.

The position is in the Management Grade. Minimal salary and allowance is Rs.1.00 lakh per annum. Housing and transport provided.

QUALITY CONTROL MANAGER

M.Sc./Ph.D. in Organic or Analytical Chemistry with minimum 5 years of post qualification experience in the development of Analytical Test methods or Quality Assurance. He should have hands-on familiarity with HPLC, GLC and other instrumental methods of analysis. He should be able to lead a team of technicians and will be responsible for releasing materials required for production and despatch. A good working knowledge of GLP and ISO-9000 validation criteria is preferred. Excellent oral and written communication skills and the ability to work as part of a team are essential.

The position is in the Management Grade. Minimal salary and allowance is Rs.1.00 lakh per annum. Housing and transport provided.

RESEARCH OFFICERS

M.Sc. in Organic or Analytical Chemistry with minimum two years post qualification experience in Organic synthesis, Quality Control or Instrumentation. They are expected to develop synthetic and analytical methods for various speciality chemicals under the guidance of Research scientists. Starting gross salary is Rs.5000/- per month.

RESEARCH ASSISTANTS/TRAINEES

B.Sc./M.Sc. with consistently good academic record with or without research experience. Fresh candidates will be absorbed as trainees on a monthly stipend of Rs.2000/- to Rs.3000/-.

THERE IS NO LAST DATE FOR APPLICATION. EXCELLENT GROWTH PROSPECTS FOR BRIGHT HARD WORKING CANDIDATES.

HIGHER STARTING SALARY FOR DESERVING CANDIDATES.

Apply to:- Director —
EPIC Enzymes, Pharmaceuticals & Industrial Chemicals Ltd.
514, Persepolis, Sector 17, Vashi, New Bombay 400 705

FERRIERA ASSOC./EPIC/1244/95

CURRENT SCIENCE

Volume 69 Number 1 10 July 1995

CORRESPONDENCE

Fancy and facts in science and art

Science is often defined as the search for truth. In a true sense, art can also be considered as the search for truth. This leads us to the often-debated fundamental question – What is truth (*'Satyam'*)? *'Sat'* in the Sanskrit word *'Satyam'* means 'that which exists forever'. This eternal fact was aptly presented by our predecessors as – *'Ekam sat vipra bahudhavadanti'* (Truth is one; scholars present it in many ways).

It is, therefore, not uncommon to feel the intermingling of the scientific quest for precision with artistic values of aesthetics. History does provide many inspiring examples. I was impressed by a statement from Charles Lindberg, whose exploits with aircrafts are legendary: 'Science, freedom, beauty, adventure: What more could you ask of life? Aviation combined all the elements I loved. There was science in each curve of an airfoil, in each angle between strut and wire, in the gap of a spark plug or the colour of the exhaust flame. There was freedom in the unlimited horizon, on the open fields where one landed. A pilot was surrounded by beauty of earth and sky. He brushed treetops with the birds, leapt valleys and rivers, explored the cloud canyons he had gazed at as a child. Adventure lay in each puff of wind'.

Barbara McClintock, undoubtedly one of this century's outstanding scientists, is a shining example whose work has embodied the elements of beauty and adventure. She dared to think differently and firmly placed her work and vision on genetic instability, duly earning a Nobel prize in 1983, though it is almost 34 years late! Aren't the anthocyanin colour patterns on the maize kernels, which incited the interest of McClintock, aesthetically appealing and beautiful in their own right?!

Probably, every scientist who takes pride in his/her work consciously realizes the innate artistic beauty of the subjects. I was, therefore, not so surprised to read the following statement by Chandrashekaran[1] in an article on biological research: 'What under the sun can possibly surpass the beauty of the feedback physiological principles involved in the capture of a moth by an insectivorous bat in the pitch darkness of a new moonlight?'

Doesn't art involve scientific elements? What better example can there be but the graphic work of M. C. Escher? Although he declared himself as 'absolutely innocent of training or knowledge in the exact sciences', Escher's drawings in the early part of this century gave expression to abstract concepts of infinity from finity, symmetry, self-similarity, precision and periodicity, stimulating the interest of mathematicians and crystallographers[2]. Aren't all these elements reflected in a good measure in the biological world as 'pattern formation', whether it is the colour patterns in the butterfly wings, anthocyanin patterns on a maize kernel, spatial periodicity of plant organs or the exine architecture of pollen?! So much so, Nijhout[3], a pioneer researcher on wing colour patterns, remarks: 'Few things in nature match the beauty and variety of the patterns on the wings of the butterflies and moths.'

Although there are several unifying principles in science and art, an important point to consider, in this context, is the degree and freedom of 'creativity'. I do find (as many must have) a basic difference in the 'creative potential' of science and art. Although science does give every scientist unlimited opportunities (at least theoretically) for demonstrating his own ego, talent, imagination and industry as in arts, the creative potential and the freedom of imagination of a poet or a writer are indeed far greater than that of a scientist. A poet could thus grant imaginative boons to a *'rakshasa'* (Hiranyakaship) that he could neither be killed by man nor by an animal, neither by a living nor by a dead weapon, neither in the day nor at night, neither on the earth nor in the sky and neither in the house nor outside. Having set all these variables, he comes out with a wonderful answer by creating a marvel of imagination – *'Narasimha'* (the Lion God) and the rest is an awe-inspiring mythological story. It is, therefore, said: *'Apaare kavyasamsare kavireva prajapathihi; ethamsmyi rochite viswam tathedham parivartate'*, broadly meaning that poet is a 'brahma' of the vast literary family; as he wishes, so he creates.

The imagination and the infinite expressions from a finite means thus depend on the talent of an artist, whether in music, painting, poetry or even in film-making, where a Spielberg can recreate dinosaurs (aided by computer technology, of course), give 'life ' to these and instantly appeal to a worldwide audience.

The ability of a scientist in playing the role of a 'creator' has increased manifold in the recent years in the biological field after the advent of genetic engineering. For instance, let us consider the recent manipulations of floral colour patterns through molecular techniques, termed 'molecular flower breeding'[4]. In nature, we do not find brick-red petunias or blue roses for the simple reason that these plants lack appropriate enzymes that catalyse certain chemical reactions in the anthocyanin biosynthesis to produce respective colour pigments. Researchers have been recently successful in introducing and expressing the gene coding for the enzyme dihydroquercetin reductase from maize into petunia, thus producing a novel brick-red petunia variety. This is

a fairly successful, commercially beneficial and environmentally harmless (apparently) application of plant genetic engineering. In contrast, several restrictions do exist in carrying out similar exercises in the animal world. The point I make is that a scientist cannot and perhaps should not have a free hand in his 'creations' as much as an artist has, and his pursuits are most often, if not always, limited by nature and other regulations imposed by man himself. Except for this variable degree of 'freedom', there may not be any fundamental difference between scientific and artistic quests for truth. The underlying similarities between science and art have been presented in a brilliant portrayal of Newton and Michelangelo by Prof. Chandrasekhar[5].

Finally, it is tempting to quote Leonardo da Vinci (cited by Dobzhansky[6]), the multifaceted personality par excellence: 'Even though the genius of man might make various inventions, attaining the same end by various means, it will not invent anything more beautiful or more economical or more direct than nature, for in nature's inventions nothing is wanting and nothing is superfluous.'

1. Chandrashekaran, M. K., *Curr. Sci.*, 1991, **61**, 309–311.
2. Schattschneider, D., *Sci. Am.*, 1994, **271**, 48–53.
3. Nijhout, H. F., *Sci. Am.*, 1981, **245**, 104–115.
4. Mol, J., Stuitje, A., Gerats, A., van der Krool, A. and Jorgenssen, R., *Trends Biotech.*, 1989, **7**, 148–153.
5. Chandrasekhar, S., *Curr. Sci.*, 1994, **67**, 497–499.
6. Dobzhansky, T., *Cold Spring Harbor Symp. Quant. Biol.*, 1959, **24**, 15–30.

B. M. PRASANNA
Division of Genetics
Indian Agricultural Research Institute
New Delhi 110 012, India

Distortion of Indian degrees

It is rather common to find that authors of research papers who are of Indian origin but now settled abroad mention their Indian degrees in a distorted form in their biographies. Here are some examples from recent IEEE publications:

1. '... received the B Eng from the Indian Institute of Science, Bangalore, in 1964 ...'
2. '... received the B S E E degree from the Indian Institute of Science in 1976 ...'
3. '... received the B S E E degree in Electrical Engineering from Bangalore University, India, in 1969 ..'
4. '... received the B S degree in electronics from the Bangalore University in 1981 ...'
5. '... received the B Eng (Electronics and Telecommunication Engineering) degree from the Bengal Engineering College, University of Calcutta in 1980 ...'
6. '... received the B S degree from Agra University, India in 1969 and M S degree from Aligarh University, India ...'
7. '... received a B S (Honours) degree in Physics from the University of Delhi, St. Stephens's College, in 1956 ...'
8. '... received the B S in Electrical Engineering from the University of Madras, India ...'

The names and references have been suppressed for obvious reasons. All the degrees mentioned above are distorted versions, as one may easily verify. IISc neither offers a B Eng nor a B S E E; neither does Bangalore University award a B S E E or B S in Electrical Engineering or a B S degree in Electronics. Calcutta University awards a B E and not a B Eng degree. None of the three Universities at Agra, Delhi or Madras offers a B S degree. Also, Aligarh University does not award an M S degree.

Why do these authors, who are some of our very best products, distort their Indian degrees? Is it their wish to conform to the US/Canadian degree? If so, why? If conformity is important, why is it that scientists of European origin do not have to resort to such distortions? Or, is it a reflection of an inferiority complex? Are B Sc, M Sc and B E inferior nomenclatures as compared to B S, M S and B S E E?

It is my impression that Indian degrees are well understood and well respected throughout the world. Such disrespect for Indian degrees, as shown by some authors, is, in my opinion, totally uncalled for, and almost amounts to falsification. What would happen to the individual concerned if one of our institutions writes to his/her employer that he/she was never awarded the stated degree?

S. C. DUTTA ROY
Electrical Engineering Department
Indian Institute of Technology
New Delhi 110 016, India

On 17 April 1995 the National Working Group on Patent Laws (NWGPL) – a nationwide non-government organization active in the area of Intellectual Property Rights (IPRs) – issued an appeal for preservation of the essential elements of the Indian Patents Act, 1970. Although it is no one's case that changed circumstances do not call for reappraisals, the scientific community has, so far, been admitted to only the outer periphery of the present debate and national decision-making on IPRs. This is all the more surprising because the generators, preservers, advancers, users, transmitters and first owners of the products of the intellect – intellectual property – are, in large measure, scientists and engineers, working within, and receiving approbation and remuneration from, the larger polity.

— Editors

Ammending the Patents Act – An appeal from the National Working Group on Patent Laws

In its appeal, the National Working Group on Patent Laws reiterated that what is at stake is not a narrow technical or legal issue. 'We are witnessing nothing short of an ever-expanding annexation of our sovereign spaces of vital decision-making. The proposed amendment of the Patents Act is being imposed on the nation. ... in line with the exclusionist and inequitable approach of the industrialized countries and their transnational corporations. The TRIPs agreement is nothing but a comprehensive and meticulous embodiment of that approach. ... [the] proposed amendment to the Indian Patents Act of 1970 [is] to make it consistent with the TRIPs agreement.'

'The issue is political and transcends "technical" or "legal" considerations. It is therefore, inappropriate to think of introducing "safeguards" in the proposed amendment. ... For the same reason, it is misleading to confine one's outlook to sectoral or superficial adverse effects and then to plead that there will be no real damage for say "seven" years.'

'There is also no question of our country ceasing to be a member of the World Trade Organization (WTO) automatically if the proposed amendment falls through [in Parliament]. Leaving aside the question of whether it is the national interest to be member of the WTO, with all its detrimental contents and its continuing, intrusive legislative mandate overriding the sovereignty of its members, there is no way of our being expelled from the WTO, only because our sovereign Parliament has decided not to approve the amendment. In many international agreements, the US executive committed itself to the requirements of these agreements after thorough negotiations, but such commitments were rejected or modified by the US legislature in the exercise of its sovereign power. The result was that the agreements were amended to accommodate US and *not* that the US was expelled from the agreements. Moreover, the expulsion from the WTO requires a decision by a three-fourths majority of its members. Such mobilization in WTO against India on this issue is next to impossible. ...'

'An apprehension is voiced that unilateral punitive action may follow. First, for any WTO-legal action, the member concerned will have to follow the route of dispute settlement procedures, which can be activated only if there is a specific cause of action. Moreover, the normal working of a democratic process can hardly provide such a cause of action for other members. However, this does not rule out unilateral, WTO-illegal action on the part of any member who wishes to bully us. But *ab initio* such action is WTO-illegal. And all the so-called protection of a rule-based system then becomes a mirage.'

A detailed analysis by the NWGPL of the provisions of the Patents (Amendment) Bill, 1995, follows.

S. no	Provision in the Patents (Amendment) Bill, 1995	Remarks
1	Preamble Whereas India is a signatory to the Agreement for the establishment of the World Trade Organization including the Agreement on Trade Related Aspects of Intellectual Property Rights for the purpose of reduction of distortions and impediments to international trade and promotion of effective and adequate protection of intellectual property rights.	There is nothing to show that the Indian Patents Act, 1970, creates distortions and impediments to trade. In the light of this it would not be desirable to resort to self-condemnation of our Patent Laws.
2	Section 24A(1) Notwithstanding anything contained in sub-section (1) of Section 12, the controller shall not, under that sub-section, refer an application in respect of a claim for a patent covered under sub-section (2) of Section 5 to an examiner for making a report till the 31st day of December 2004 and shall, where an application for grant of exclusive right to sell or distribute the article or substance in India has been made in the prescribed form and manner and on payment of prescribed fee, refer the application for patent, to an examiner for making a report to him as to whether the invention is not an invention within the meaning of this Act in terms of Section 3 or the invention is an invention for which no patent can be granted in terms of Section 4.	It would not be desirable to keep the patent application unattended. It is important to examine the patent applications in accordance with Section 12 of the principal Act, in addition to Sections 3 and 4. The TRIPs Text specifies that exclusive marketing right is to be given and not exclusive right to 'sell or distribute'. Exclusive right for distribution may not be tenable under the MRTP Act. It is important to examine whether the article or substance is not in public domain. The invention has to be new, involve an inventive step and should be capable of industrial application to qualify for an exclusive marketing right.

3 Section 24A(2)

Where the Controller, on receipt of a report under sub-section (1) and after such other investigation as he may deem necessary, is satisfied that the invention is not an invention within the meaning of this Act in terms of Section 3 or the invention is an invention for which no patent can be granted in terms of Section 4, he shall reject the application for exclusive right to sell or distribute the article or substance.

as above

4 Section 24A(3)

In a case where an application for exclusive right to sell or distribute an article or a substance is not rejected by the Controller on receipt of a report under sub-section (1) and after such other investigation, if any, made by him, he may proceed to grant exclusive right to sell or distribute the article or substance in the manner provided in Section 24B.

In the TRIPs Text, the connotation used is exclusive marketing right and not exclusive right to sell or distribute. Public objection should also be invited.

5 Section 24B(1)

Where a claim for patent covered under sub-section (2) of Section 5 has been made and the applicant has:
(a) where an invention has been made whether in India or in a country other than India and before filing such a claim filed an application for the same invention claiming identical article or substance in a convention country on or after the 1st day of January, 1995 and the patent and approval to sell or distribute the article or substance on the basis of appropriate tests conducted on or after the 1st day of January 1995, in that country has been granted on or after the date of making a claim for patent covered under sub-section (2) or Section 5; or

(i) Patent must be obtained in the country of invention on a product prior to the claim made under sub-section (2) of Section 5.

(ii) If the patent application has been rejected in other convention countries, it would be desirable to know the status of such applications.

6 Section 24B(1) (b)

Where an invention has been made in India and before filing such a claim, made a claim for patent on or after the 1st day of January 1995 for method or process of manufacture for that invention relating to identical article or substance and has been granted in India the patent therefor on or after the date of making a claim for patent covered under sub-section (2) of Section 5,

Process patents are already granted in India and therefore EMR on products arising out of patented processes cannot be granted.

7 Section 24B(1) (b) – sub-para

and has received the approval to sell or distribute the article or substance from the authority specified in this behalf by the Central Government, then he shall have the exclusive right by himself, his agents or licensees to sell or distribute in India the article or the substance on and from the date of approval granted by the Controller in this behalf till a period of five years or till the date of grant of patent or the date of rejection of application for the grant of patent, whichever is earlier.

(i) The relevant authority, e.g., Drug Controller, from whom marketing approval is to be taken should specifically be mentioned in the text.

(ii) Public opposition to marketing approval should be invited.

8 Section 24B(2)

Where the specifications of an invention relatable to an article or a substance covered under sub-section (2) of Section 5 have been recorded in a document or the invention has been tried or used, or the article or the substance has been sold, by a person, before a claim for a patent of that invention is made in India or in a convention country, then the sale or distribution of the article or substance by such a person, after the claim referred to above is made, shall not be deemed to be an infringement of exclusive right to sell or distribute under sub-section (1).

Since the article or the substance is already in the public domain, no exclusive right can be given.

9 Section 24C

The provision in relation to compulsory licences in Chapter XVI shall, subject to the necessary modifications, apply in relation to an exclusive right to sell or distribute under Section 24B as they apply to and in relation to a right under a patent construed accordingly, namely: to sell or distribute and for that purpose the following modifications shall be deemed to have been made to the provisions of that chapter and all their grammatical variations and cognate expressions shall be construed accordingly, namely:

(i) Indian Patent Act, 1970, provides for Licences of Right which is more appropriate as compared to compulsory licences.

(ii) Licences of Right should only be granted to encourage local production and not marketing.

(a) throughout Chapter XVI,
  (i) working of the invention shall be deemed to be selling or distributing of the article or substance,
  (ii) references to 'patents' shall be deemed to be references to 'right to sell or distribute',
  (iii) references to 'patented article' shall be deemed to be references to an 'article for which exclusive right to sell or distribute has been granted'.

Local production, and not 'selling and distribution', should be considered as working.

Binding and release of polypeptides during groEL-mediated folding: Its cellular significance?

Amitabha Chaudhuri

groEL is an *E. coli* homologue of the hsp-60 family of chaperone proteins called chaperonins. Besides groEL, the chaperonin family of proteins includes rubisco subunit-binding protein and hsp-60, which share 60% amino acid identity among themselves[1,2]. The eukaryotic protein t-complex polypeptide-1 (TCP1), although unrelated to hsp-60 class of proteins, is also a member of the chaperonin family[3].

The groEL protein was originally identified in *E. coli* as a component whose deficiency resulted in defective morphogenesis of bacteriophage λ, T4 head structures and T5 tail structures[4,5], thus suggesting for it a role in protein assembly. The gene encoding the groEL protein, groE, is essential for bacterial growth at both normal and heat shock temperatures[6]. During heat shock, the groEL protein increases 10-fold, thereby constituting 10% of the total cytoplasmic proteins[7]. The induction of groEL under stress suggests that the function of the protein is required to a greater extent under stressful conditions, when cellular proteins are likely to be destabilized with the loss of their native structure.

Evidence for the participation of groEL in protein folding comes from experiments done both *in vivo* and *in vitro*. Mutants of groEL failed to promote proper folding of their cytoplasmic proteins[8], suggesting that groEL assists newly synthesized cytoplasmic proteins to reach their native tertiary structure. groEL also promotes oligomerization of rubisco[9] and folding of a host of other unfolded *E. coli* proteins *in vitro*[10].

groEL is a large cylindrical oligomeric protein with 14 subunits arranged as two-stacked rings of seven subunits each[11,12]. Each ring forms a cage-like structure[13] which binds polypeptides in a central cavity[14] (Figure 1). Cryoelectron microscopy of groEL and complexes of groEL with ATP and groES have revealed conformational changes in the molecule with the formation of each complex[15]. Each subunit of groEL probably has an ATP-binding site[16]. The stoichiometry of substrate binding indicates that the 840 kDa groEL particle binds one or two substrate molecules. The substrate protein is believed to be sequestered in a 6 nm central cavity as a compact-folding intermediate[17].

The crystal structure of groEL at 2.3 Å resolution[18] has thrown light on the structure–function relationship of the chaperonin protein[19]. Each ring of seven subunits is folded into three distinct domains: the apical domain, which binds the substrate polypeptide; the equatorial domain, that forms the contact surface between the two rings; and the intermediate domain that connects the apical to the equatorial domain. The latter harbours the ATP-binding site.

The structure of groEL has been used as a framework for an extensive mutational analysis by Fenton *et al.*[20]. Mutant groEL molecules defective in the binding of one or more of its three ligands – unfolded polypeptide, groES and ATP – have been obtained. The polypeptide-binding domain mapped exclusively to that segment of the apical domain that faces the central channel. In this domain there are eight highly conserved hydrophobic amino acids, and changing any one of them is sufficient to abolish polypeptide binding. The amino acid composition of the polypeptide-binding domain confirms the prediction that the chaperonin protein interacts with the substrate polypeptide through hydrophobic patches.

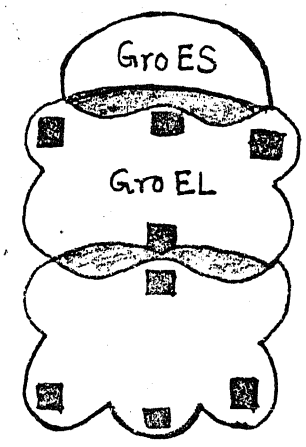


Figure 1. A diagrammatic representation of groEL–groES–ATP complex. groEL is shown as two-stacked rings with bound ATP (■). The shaded region represents the cavity within the groEL cylinder in which the substrate polypeptide is sequestered.

The same mutations that abolish polypeptide binding also affected the binding of groES. This raises the interesting possibility that the release mechanism of the polypeptide may involve a physical displacement by groES. The crystal structure of groEL and the mutational analysis has, therefore, vindicated many of the predictions that were derived from biochemical evidences.

Contribution of groEL in protein folding

The classic *in vitro* studies of Anfinsen[21] showed that folding of proteins is apparently a spontaneous process solely dictated by the primary structures. This conclusion seems now to be an oversimplification owing to markedly different conditions of protein folding *in vitro* and *in vivo*.

Whereas *in vitro* folding reactions are initiated at low temperatures, at low protein concentrations and with a complete polypeptide chain, *in vivo* folding occurs at high temperatures, at high protein concentrations and even before complete synthesis of polypeptide has occurred. As a consequence of the conditions *in vivo*, a nascent unfolded polypeptide chain is likely to interact with other unfolded protein molecules through their interactive surfaces, forming aggregates. The role of the chaperonin protein is to prevent unproductive incorrect interactions between these polypeptides which would otherwise act as kinetic dead-end traps[3]. This function is accomplished by noncovalent binding of groEL to exposed interactive surfaces of proteins, producing stable complexes. groEL might also interact with proteins trapped in misfolded forms, enabling them to assume productive correctly folded conforma-

tions[22]. These functions of the chaperonin protein require that they interact with the substrate proteins promiscuously and also bind unfolded proteins with higher affinity than their native counterparts.

groEL binds to a large fraction of denatured *E. coli* proteins[10]. Using labelled pre-β-lactamase and chloramphenicol acetyl transferase it was shown that upon their synthesis in a bacterial translation system, these proteins form complexes with groEL. The bound proteins could be displaced only by denatured myoglobin, indicating affinity of groEL for denatured substrates[23]. groEL also prevents proteins from undergoing aggregation upon dilution from a denaturant. Using rhodanese and dihydrofolate reductase it was shown that both the proteins bound groEL in a state intermediate between unfolded and the native state as characterized by proteinase sensitivity[17], affinity to ANS and intrinsic fluorescence properties[16, 24]. Thus, groEL provides a microenvironment where unfolded proteins are kept isolated from the bulk milieu, preventing unproductive interaction and thereby enabling proper folding.

The binding of unfolded proteins to groEL does not assure their proper folding. Complete folding requires the participation of a co-chaperonin groES, and ATP hydrolysis[25, 26]. The role of groES is not well-characterized except that it increases the cooperativity of binding and hydrolysis of ATP by groEL[27].

The hydrolysis of ATP during groEL-mediated folding suggests that unlike protein folding *in vitro*, protein folding *in vivo* is energy-consuming. This energy barrier is likely to serve three functions: (a) preventing free exchange of groEL-bound proteins with the other unfolded polypeptides; (b) allowing limited folding of the polypeptide on the surface of groEL and (c) facilitating the release of the protein from groEL.

The release of the protein from the chaperonin protein raises the question of whether the protein is released in its native or nonnative conformation. If a protein is released in its native conformation, it will not interact further with the chaperonin protein. However, if the protein is released in a nonnative conformation, it is likely to interact with other groEL molecules till it reaches the native state. Therefore, groEL-mediated folding of proteins involves either one or several rounds of binding and release of the polypeptides. How does one distinguish between these possibilities?

This question has been addressed rather elegantly in a paper by Weissman *et al.*[28], who have developed a technique to capture partially folded substrate proteins should they be released intermittently during their folding in the presence of groEL. In this technique mutant groEL molecules are employed, which bind substrate proteins only in their unfolded/partially folded form but do not release them even in the presence of Mg-ATP. Using ^{35}S-labelled rhodanese and ornithine transcarbamylase, they have shown that when mutant groEL protein is added in a folding reaction initiated with wild-type groEL, groES and Mg-ATP, it results in the capture of labelled proteins by the mutant groEL. In the absence of either groES or Mg-ATP, the mutant groEL fails to trap the enzymes. This observation suggests that the substrate protein is indeed released from groEL during folding and this release requires the participation of groES and ATP hydrolysis. A kinetic analysis of the process revealed that whereas complete folding requires 7 min (judged by the recovery of enzyme activity), the mutant groEL was able to trap the folding intermediates within 0.5 min of initiating the folding reaction, indicating that the substrate protein was unable to attain its native conformation upon interaction with one chaperonin molecule. To determine whether the process of binding and release of polypeptides occurs throughout the folding reaction, mutant groEL was added at different time intervals after the folding reaction was initiated. Addition of mutant groEL to the folding reaction 2 min after initiation resulted in the capture of more than 60% of the substrate polypeptides. When the mutant groEL was added 5 min after the initiation of the folding reaction, about 50% of the substrate protein was captured. This observation of Weissman *et al.*[28] suggests that the release of the polypeptide occurred throughout the folding reaction, and with increased time interval the amount of unfolded protein in the reaction mixture declined. Therefore, chaperonin-mediated folding of proteins involves several rounds of binding and release from groEL. Each cycle of binding and release is expected to require hydrolysis of 14 ATP molecules, suggesting that the requirement of 100 ATP molecules for the *in vitro* folding of one rhodanese molecule may not be an overestimation[16].

Why cannot a protein complete its folding without going through iterative cycles of binding and release? It could save the cell a substantial amount of energy. The fact that this process has been selected for in spite of it being energy-inefficient suggests that there may be something important associated with this process such that the cycles of binding and release become indispensable. The iterative interaction of a polypeptide with groEL may help to serve the following important functions:

1. As mentioned earlier, the host polypeptide is noncovalently associated with groEL through its interactive surfaces. These surfaces are, therefore, unavailable for any intramolecular interaction during the folding of the polypeptide. This is likely to result in an impasse in which the release of these interactive surfaces from groEL becomes a prerequisite for folding to proceed. In such a situation, the partially folded polypeptide has two options: (i) it can fold within the cavity of groEL by sequential release of the interactive surfaces from one or more subunits of the groEL molecule, as has been suggested by Hendrick and Hartl[17]; (ii) or it can get released from the chaperonin protein into the bulk milieu and undergo intramolecular interactions to reach their native state. The first option to be fruitful requires that the noncovalently associated interactive surfaces of the polypeptide be released from groEL in a sequential manner as defined by the folding pathway of the protein. Since this is difficult to achieve mechanistically, the second option of dissolving all interactions simultaneously and releasing the polypeptide in the bulk solution is expected to occur.

2. In the assembly of oligomeric proteins individual subunits interact with each other to form the oligomer. The interaction between subunits during oligomerization requires that each subunit bear specific patches on themselves that are recognized by other subunits. The presence of such exposed interactive patches is likely to promote aggregation, especially if the subunits are allowed to interact with each other randomly. The role of the chaperonin protein is, therefore, to hold the monomers, thereby preventing any unproductive interaction, and release them at a time when fruitful interaction

will result in proper oligomerization.

3. Proteins with sulphydryl groups or proline residues undergo thiol/disulphide exchange reaction and proline *cis/trans* isomerization during folding[29, 30]. These reactions are catalysed by the enzymes protein disulphide isomerase and proline *cis/trans* isomerase, respectively. Release of partially folded polypeptide from groEL is a prerequisite for it to interact with these enzymes.

4. Proteins may be marked for degradation during rounds of binding and release from the chaperonin molecule. The participation of groEL and groES in the breakdown of CRAG (a fusion protein containing at its amino terminus 12 amino acids of the cro repressor, followed by a truncated form of protein A and 14 amino acids of β-galactosidase at its carboxy terminus) has been demonstrated[31]. Although the mechanism is not understood, association of CRAG with groEL seems to be the rate-limiting step in CRAG degradation. This observation reveals yet another complexity in the function of groEL.

5. The release of a polypeptide from groEL may be important in limiting the amount of time a polypeptide may spend on the chaperonin molecule. This is likely to prevent the chaperonin proteins from remaining locked up in a bound state for a long time, thereby making them unavailable for other unfolded proteins. This is expected to assume importance during stressful conditions, under which the unfolded proteins are likely to outnumber the chaperonin molecule. In the groEL-mediated folding, the release of polypeptides, therefore, assumes importance in defining the number of cycles of interactions each polypeptide undergoes before reaching its native state. In other words, some proteins may require one round of binding and release while other proteins may require many before reaching their respective folded conformations.

The requirement of groEL in assisting the folding of a host of proteins both *in vivo* and *in vitro* raises the question whether all proteins require chaperonins for folding *in vivo*? If not, how do chaperonins distinguish between these proteins, especially when every protein molecule starts its life from an unfolded polypeptide chain? One explanation can be that a protein which does not require a chaperonin for folding *in vivo* has less interactive surfaces for the binding of the chaperonin protein, or the interactive surfaces may become inaccessible during early steps of folding. These proteins, therefore, not only obviate the need of a chaperonin but also avoid unproductive interactions with other protein molecules during the course of folding. In this light, one can, in fact, classify proteins as being slow- or fast-folding, and in line with

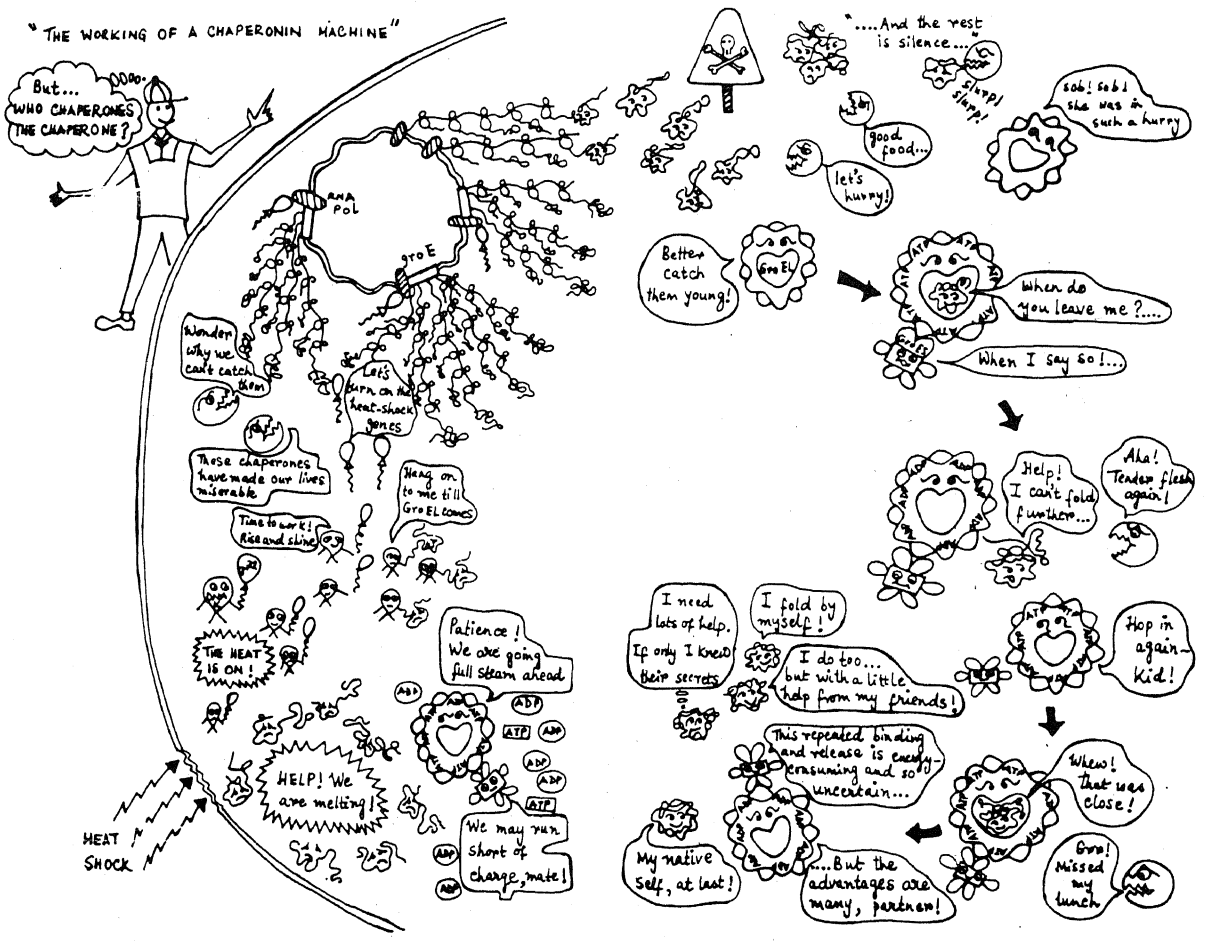


Figure 2. The working of a chaperonin machine.

this argument, the slow-folding type is likely to require the assistance of chaperonin proteins for their proper folding. The trapping method developed by Horwich's group will be invaluable in classifying proteins into these two groups. This can be achieved by determining the relative amount of time a protein spends on a chaperonin molecule before completion of chaperonin-mediated folding. For this, different proteins can be tested for their ability to fold *in vitro* in the presence of groEL. After the folding reaction is initiated with wild-type groEL, groES and Mg-ATP, the mutant groEL can be added to the reaction mixture after different periods of time and the relative amount of substrate protein that is trapped by the mutant groEL determined. A protein which goes through less number of iterative cycles of binding and release before reaching its native structure will be inefficiently trapped by the mutant groEL in contrast to proteins which complete fold through many such iterative cycles. The former can be classified as fast-folding and the latter as slow-folding types. A comparison of the sequences of these two classes of proteins is likely to reveal sequence motifs that are important in limiting the interaction of any protein with the chaperonin.

In conclusion, the detailed working of a chaperonin machine is depicted in Figure 2 with the help of a cartoon diagram.

1. Hemmingsen, S. M. *et al.*, *Nature*, 1988, **333**, 330–334.
2. Hemmingsen, S. M. *et al.*, *Nature*, 1992, **357**, 650.
3. Ellis, R. J., *Phil. Trans. R. Soc. Lond. B*, 1993, **339**, 257–261.
4. Gergopolous, C. P. *et al.*, *J. Mol. Biol.*, 1973, **76**, 45–60.
5. Sternberg, N., *J. Mol. Biol.*, 1973, **76**, 29–43.
6. Fayet, O. *et al.*, *J. Bacteriol.*, 1989, **171**, 1379–1385.
7. Neidhardt, F. C. *et al.*, *Annu. Rev. Genet.*, 1984, **18**, 295–329.
8. Horwich, A. L. *et al.*, *Cell*, 1993, **74**, 907–917.
9. Goloubinoff, P. *et al.*, *Nature*, 1989, **342**, 884–889.
10. Vittanen, P. V. *et al.*, *Protein Sci.*, 1992, **1**, 363–369.
11. Hendrix, R. W., *J. Mol. Biol.*, 1979, **129**, 375–392.
12. Carazo, J. M. *et al.*, *J. Struc. Biol.*, 1991, **106**, 211–220.
13. Saibil, H. R. *et al.*, *Curr. Biol.*, 1993, **3**, 265–273.
14. Braig, K. *et al.*, *Proc. Natl. Acad. Sci. USA*, 1993, **90**, 3978–3983.
15. Chen, S. *et al.*, *Nature*, 1994, **371**, 261–264.
16. Martin, J. *et al.*, 1991, *Nature*, 1991, **352**, 36–42.
17. Hendrick, J. P. and Hartl, F. U., *Annu. Rev. Biochem.*, 1993, **62**, 349–384.
18. Braig, K. *et al.*, *Nature*, 1994, **371**, 578–586.
19. Hartl, F. U., *Nature*, 1994, **371**, 557–559.
20. Fenton, W. A. *et al.*, *Nature*, 1994, **371**, 614–619.
21. Anfinsen, C. B., *Science*, 1973, **181**, 223–230.
22. Zahn, R. *et al.*, *Nature*, 1994, **368**, 40–45.
23. Bochkareva, E. S. *et al.*, *Nature*, 1988, **336**, 254–257.
24. Horwich, A. L. and Willson, K. R., *Phil. Trans. R. Soc. Lond. B*, 1993, **339**, 313–326.
25. Mendoza, J. A., *J. Biol. Chem.*, 1991, **266**, 13044–13049.
26. Vittanen, P. V. *et al*, *Biochemistry*, 1991, **30**, 9716–9723.
27. Zahn, R. and Plukthun, A., *Biochemistry*, 1992, **31**, 3249–3255.
28. Weissman, J. S. *et al.*, *Cell*, 1994, **78**, 693–702.
29. Freedman, R. B., *Cell*, 1989, **57**, 1069–1072.
30. Fischer, G. and Schimd, F. X., *Biochemistry*, 1990, **29**, 2205–2212.
31. Kandror, O. *et al.*, *J. Biol. Chem.*, 1994, **269**, 23575–23852.

ACKNOWLEDGEMENT. I thank Drs Kakoli Mukherjee and Anindya Sinha for sharing their views on the manuscript and helping me in drawing the cartoon diagram.

Amitabha Chaudhuri is in the Institute of Microbial Technology, P.B. No. 1304, Sector 39-A, Chandigarh 160 014, India.

COMMENTARY

Importance of basic research in the context of global change*

R. Chidambaram

This is a time of transition for Indian science. There is economic liberalization at home and there is the end of the cold war – at least in its earlier avatar – abroad. And there is a demand for environmentally benign and eco-friendly sustainable development. How should the Indian science and technology system react to these changes?

Any activity – political, military, economic, scientific or cultural – must have only one objective in view, namely, how to improve the quality of life of our citizens while ensuring the security of the country. We should quickly put India on to the fast track of national development and keep it there, without sacrificing the values for which the country has always stood for. We have the unfortunate tendency, as Indians, to denigrate ourselves too much. I do not think we have done too badly in any sphere since independence, though our talents and our resources could have made higher levels of achievements possible.

The attack on the Indian university system explicit in the proposal for the National Science University – discussed so well in recent issues of *Current Science* – is a case in point. The arrogance of that proposal is matched only by its absurdity. How can one ignore the excellent work done by a very large number of Indian scientists who have spent all the time of their professional careers in this country and imagine that a group of NRI scientists, working part-time and starting from scratch, will transform the scientific education scenario in the country? Have these people heard of

*Based on the lecture given at the Symposium on S&T Policy and Economic Reforms, organized by Dr P. S. Lokanathan Memorial Society, Bangalore, on 15 March 1995.

the green revolution and the oil exploration effort which led to the discovery and exploitation of Bombay High? Are they aware of the impact that some national laboratories like the NCL and some university departments like the UDCT of Bombay University have had on the Indian chemical industry? How have Indian scientists and technologists been able to build nuclear reactors, launch satellites or build sophisticated missile systems, using young men trained indigenously?

And these achievements have been possible in spite of technology control regimes that are operating now in strategic areas of high technology, which deny India some kinds of equipment and technologies in the name of non-proliferation and missile technology control. They also serve the interests of rich and developed countries by preventing diffusion of technologies needed for development by other countries. As long as Indian industry in non-strategic areas tends to generally use yesterday's technology, it will not be affected much by these technology control regimes. But as India becomes globally more competitive, Indian industry will find such regimes becoming operational for commercial reasons also and outside sources of technology will dry up. That is when the Indian industry will begin to feel more and more the need for indigenous R & D support.

What we are talking about above is applied research (where the motivation is to gather new knowledge with utilitarian objectives) and technology development (where the motivation is to create a product or a process). The reasons to support basic research (where the motivation is to gather new knowledge to satisfy curiosity) have to be sought elsewhere.

Support to science

The spurt in support to science, both basic and applied, in USA and in erstwhile USSR, was driven by the exigencies of the Second World War and then by the Cold War. There is perceptible change now in their approach to science – perhaps temporary – which seems to be caused by systemic boredom, accompanied by personal disinterest, with only a few young native Americans and Russians opting for a career in scientific research. In USA there seems to be a feeling 'What more can Science deliver?'. In Russia, whose economy has collapsed for the present in spite of their outstanding achievements in science and technology, the tendency could be to blame the latter for their present difficulties. The situation in India is very different.

We should not lose sight of our long-term perspective while pursuing short-term goals. In pursuing short-term and essential objectives, we should also not forget that India is *not* a small country. I am convinced that there is no frontier area of science and no area related to any high technology to which India should not commit itself – the extent of the commitment should depend on how far in the future such areas are likely to have an impact on India. To take a few examples, we should take measures to eliminate vitamin A deficiency in children and pursue other aspects of preventive health care but also participate in the global search for cures for AIDS and cancer. There is also no conflict between delivering economical non-conventional energy systems to rural areas and doing research on fast breeder reactors. India's relatively higher thrust must, of course, be in areas related to critical and strategic technologies and to agriculture, biotechnology, information technology and materials technology. But other areas must not be neglected.

Basic and applied research

The distinction between basic research and applied research is often blurred. The study of the origin of life or of the origin of the universe are fascinating fields but are clearly examples of basic research. But there are many situations where motivations for both kinds of research coexist. One of my colleagues is studying the structure of the HIV protease enzyme using the X-ray crystallography technique, which is an interesting basic research problem in itself. But his other purpose is to design an inhibitor for this enzyme which is the one which opens the protein coat of the AIDS virus and makes it infectious. My own studies on the equation of state and phase transformations at high pressures have had both a basic research motivation and an earlier applied research one related to peaceful nuclear explosion phenomenology. Similarly other studies in BARC on fission phenomena have had both basic research interest and an applied one related to the design of better and safer nuclear reactors.

Why support basic research?

The primary motivation for government or industry support to research is to help technology development for peaceful or defence purposes. But government support to basic research should have other motivations, which should be analysed in a national perspective. There is, of course, the duty of the state in any civilized country to allow intellectuals the opportunity to let their curiosity wander in trying to understand natural phenomena or ask and to answer questions of all kinds. Imagine the possible achievements of Srinivasa Ramanujan – 'the magical genius' as Kak once called him – if he had been supported by the government.

After a lecture by Michael Faraday (1791–1867) in the Royal Society, on the newly discovered phenomenon of electricity, a politician asked him: 'That was very interesting, but will it ever be used for anything?'. Faraday replied: 'Sir, one day you will tax it'. Some results of basic research do get converted into products and processes producing a great deal of wealth. But the support to basic research in India cannot be based on such possibilities.

Basic research is needed in Indian Universities to improve the quality of teaching. A teacher who is also doing research is much more in touch with the latest developments in his subject. Basic research is needed in national laboratories to support applied research and technology development and to attract the best intellects. In both the Manhattan Project and the Pokharan PNE Experiment Project, there was need for basic research scientists. The gaps in advanced applied research and technology development efforts can often be filled only by the expertise and knowledge available with basic research scientists.

International interactions and flow of information

Success in basic research at the international level is a matter of national pride. Applied research scientists and technology developers, though they create national wealth, are unfortunately less visible.

The national psyche needs eminent scientists like C. V. Raman, Srinivasa Ramanujan, M. N. Saha, S. N. Bose, Hargobind Khorana and S. Chandrasekhar to make us feel proud. But these are

basic research scientists. So was H. J. Bhabha, who transformed himself from a basic research scientist to a science manager.

There is another important aspect of basic research which is very relevant to Indian interests. The flow of information is free in the international basic research community and this is the knowledge that could form the basis for future technology development. Information flow gets restricted when you come to applied research and is totally blocked, unless you pay for intellectual property rights, once technology development is complete. In the field of crystallography, for example, with which I am familiar, a great deal of information on drug design is freely available in the basic research community and could be profitably used by the Indian pharmaceutical industry for its own technology development.

It may appear a contradiction of sorts but it is a fact that the latest technologies are needed for basic research and are, therefore, often first developed in large basic research laboratories. It is in India's interest to participate in international mega-projects in science.

Conclusion

Basic research is, of course, the fountainhead of ideas for technology development. The applied research edifice in India today is weak, except in mission-oriented agencies. Its growth in the future will be sustained by building a strong basic research foundation *now*. Otherwise, we run the risk of stifling our future S & T agenda. Of course, the S & T community must have a commitment to indigenous technology development. And the nation and its political system must have faith in the scientists. As Frederick Seitz once said: 'The advance of science requires money given with appreciation and wisdom.'

R. Chidambaram is in Anushakti Bhavan, Chhatrapati Shivaji Marg, Bombay 400 039, India.

Census of India's biodiversity: Tasks ahead

T. N. Khoshoo

There are varying and often conflicting estimates floating around nationally and internationally regarding India's biodiversity, particularly the number of animal species. Thanks to A. K. Ghosh (Director, ZSI), reliable information has been collated on animal species. Based on the current knowledge as summarized by Ghosh[1], Khoshoo[2] and Singh[3], the total number of living species identified in India so far is 126,188 (Table 1). With the publication of this table, speculations about the extent and nature of species richness in India should be set to rest till such time a formal census is undertaken.

According to the World Conservation Monitoring Centre (WCMC)[4], the total number of species described at the global level so far is 1,604,000. However, this Centre estimates that at the global level there are likely to be 17,980,000 species. i.e. about 11 times more than the presently known species. The increase is likely to be primarily from the tropics and subtropics. However, a more realistic and a working figure for species[4] at the global level is around 12,250,000.

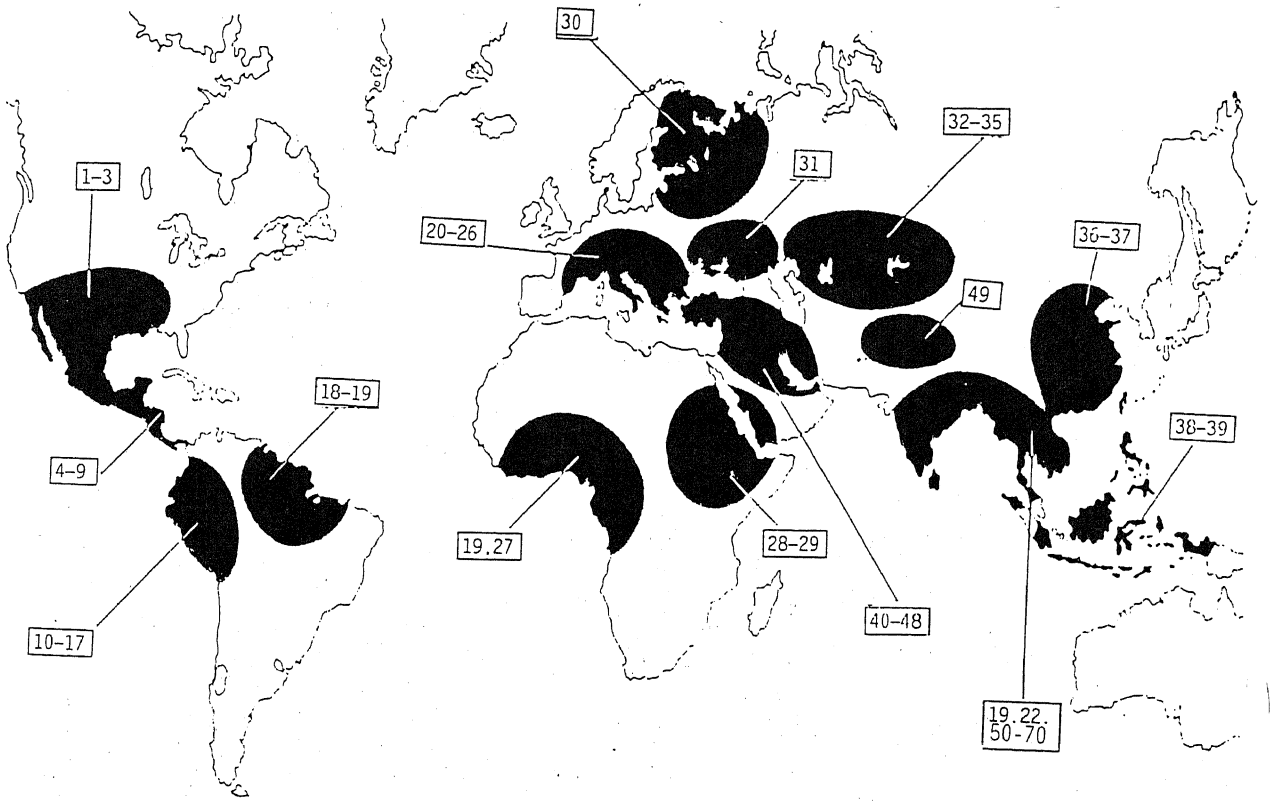
Out of the 126,188 species described from India (Table 1), Monera (bacteria) are 850 species (0.67%), Protista (Protozoa only: minus their multicellular descendants) 2577 species (2.04%), Fungi 23,000 species (18.23%), Animalia 74,875 species (59.27%) and Plantae 24,886 species (19.79%). Nearly 72% of India's biowealth is constituted by fungi (18.23%), insects (40%) and angiosperms (13.50%). This tallies generally with the overall trend seen in tropics and subtropics. Although India has only 2.4% of the land area of the world as a whole, according to the present estimates, India's contribution to the global biodiversity is around 8% species. While India stands seventh as far as the number of species contributed to agriculture (including animal husbandry) is concerned, qualita-

Table 1. Number of biota in India[1–3]

Taxon	Number of species	Percentage
Bacteria	850	0.67
Algae	2500	2.00
Fungi	23,000	18.23
Lichens	1600	1.30
Bryophyta	2700	2.14
Pteridophyta	1022	0.80
Gymnosperms	64	0.05
Angiosperms	17,000	13.50
Protozoa	2577	2.04
Mollusca	5042	4.00
Crustacea	2970	2.35
Insecta	50,717	40.00
Other invertebrates including hemichordata	11,252	9.00
Protochordata	116	0.10
Pisces	2546	2.02
Amphibia	204	0.16
Reptilia	428	0.34
Aves	1228	1.00
Mammalia	372	0.30
Total	126,188	100.00

tively India's contribution to agribiodiversity is indeed very impressive[5]. Among other species, Indian region has contributed zebu, mithun, chicken, water buffalo, camel, rice, sugarcane, banana, tea, mango, cucumber, egg plant, citrus, jute, minor millets, Asiatic vignas, brassicas, rice bean, tree cotton, jackfruit, edible dioscoreas, alocasia, colocasia, amorphophallus, cardamom, black pepper, ginger, turmeric, vegetable amaranths, cucurbits, several umbellifers, bamboos, a number of medicinals, aromatics and ornamentals, etc. (Figure 1; refs. 5 and 6). Besides, India has also been a secondary centre of domestication for horse, goat, sheep, cattle, yak, donkey, tobacco, potato, red pepper, grain amaranths, maize, soyabean, oil palm, etc. Some of these species are very early introductions into India. These have undergone evolution in the Indian region and have been selected in historical time to meet the human needs suiting Indian conditions. India has also contributed very important genes to world agriculture and animal husbandry which have made significant difference to better the human welfare[2].

All the present-day crops and domesticated animals known today were spotted and initially domesticated by the so-called early humans (particularly women). There have not been any additions to the traditional crop and domesticated animal species for the last 10,000 years. This shows the intuitive power of the ancient agriculturists for picking the right kind of animals and plants for domestication. Thereafter, modern science of genetics and breeding has not added any new crop or domesticated animal, but has no doubt increased their productivity beyond measure. However, in time to come, new additions are likely to be from amongst microorganisms for providing new products. They are likely to be the future biofactories as is the case for human insulin. Microorganisms have a big arsenal of enzymes which can perform varied and often weird tasks. Furthermore, these organisms have been performing a unique function from the day plants began to first emerge from sea and colonize totally sterile and barren land devoid of any organic matter. This was rendered possible on account of the mutualistic association between plants on the one hand, and bacteria and fungi on the other, thus helping in the uptake of nutrients by plants. Such mutualistic associations between land plants and microorganisms is seen in most of the fossil species of the then land plants colonizing the earth



1 Turkey	11 Guinea Pig	21 Cattle	31 Horse	41 Dates	51 Mithun	61 Citrus
2 Sunflower	12 Alpaca	22 Pig	32 Bactrian camel	42 Onion	52 Chicken	62 Jute
3 Tepary bean	13 Cotton	23 Grapes	33 Alfalfa	43 Peas	53 Waterbuffalo	63 Jackfruit
4 Avocado	14 Lime bean	24 Barley	34 Millet	44 Wheat	54 Banana	64 Cardamom
5 Cocoa	15 Peanuts	25 Olive	35 Hemp	45 Ass	55 Rice	65 Black pepper
6 Sweet potato	16 Pepper	26 Rye	36 Foxtail millet	46 Dromedary	56 Tea	66 Ginger
7 Maize	17 Potato	27 Watermelon	37 Soya bean	47 Sheep	57 Sugarcane	67 Turmeric
8 Runner bean	18 Pineapple	28 Finger millet	38 Coconut	48 Goats	58 Mango	68 Minor millets
9 Tomato	19 Yam	29 Sorghum	39 Breadfruit	49 Yak	59 Cucumber	69 Veg.amaranths
10 Llama	20 Goose	30 Reindeer	40 Barley	50 Zebu	60 Egg Plant	70 Asiatic Vignas

Figure 1. Centres of origin and domestication of agribiodiversity (after Boyden[6]; modified).

nearly 408 to 387 million years ago. For instance, microscopic examination of most fossil plants shows ectomycorrhiza and endomycorrhiza.

Taxonomists: A vanishing tribe

At present good taxonomists are indeed highly endangered specialists among biologists, although India is among the very few countries to have Botanical and Zoological Surveys (BSI and ZSI) which have been rendering yeoman service against all odds. These Surveys in particular and taxonomists in general are today overshadowed by the so-called and more often second-rate biotechnologists and environmentalists. There is, however, no doubt that more investments (accompanied by strict accountability) have to be made in genetics, breeding, biotechnology, and pharmaceutical sciences for gene and drug prospecting and product development. But the country has to realize that *all wisdom begins by calling all living (including humans) and nonliving things by their proper names*. This makes the role of taxonomists most important for all development involving biologicals.

Up to early 1980s, Indians have made world-class contributions in modern taxonomy, biosystematics, experimental taxonomy, circumscription of gene pools and the allied applied aspects. From this work a large number of products (by way of new cultivars) were developed which helped the country in food self-sufficiency and better varieties in horticulture and floriculture and economic animals. Such contributions emerged from some universities in general, but principally from some of the laboratories under the Indian Council of Agricultural Research (ICAR) and the National Botanical Research Institute under the aegies of the Council of Scientific and Industrial Research (CSIR).

Apart from taxonomists, a large number of persons in different trades use taxonomic and floristic information in their work. These are: teachers from primary to adult, to college and to university level for purposes of education in biology, land-use managers, environmental surveyors, geneticists and breeders, biotechnologists, biochemists, microbiologists, pathologists, anatomists, palynologists, embryologists, physiologists, pollination biologists, seed laboratories, nurserymen, forest managers, silviculturists, weed and pest control specialists, fertilizer manufacturers, food quality controllers, agricultural consultants and extension workers, animal feed companies, indigenous medical herbalists using natural medicine, pharmacognostical and pharmaceutical researchers, drug companies, natural dye specialists, dermatologists, toxicologists, allergy specialists, forensic science specialists, ecologists, environmentalists, mining engineers, real-estate appraisers, municipal and regional planners, recreation planners and managers, sports-field managers, landscapists, architects, civil engineers, interior designers, water-way managers, flood control engineers, craftsmen, custom officials, border and enforcement agencies, lawyers, mathematicians, computer scientists and a host of other specialists[7].

The kind of information needed for basic biological, agricultural, biotechnological and pharmaceutical research is: correct identification and accurate naming, reference, systematic relationship, information on intraspecific variation, cultivars, hybrids, size of population, host–parasite relationship, phenology, economic factors, variation, distribution, habitat preference, protection status, etc. This involves many disciplines like anatomy, chemistry, embryology, palynology, cytology/cytogenetics, crossability, genetics, ecology, evolutionary biology, molecular biology, population biology, size of gene pool, etc. Such information need not be generated for all the species, but for those that are related to agribiodiversity and drug industry. Furthermore, conservation biology is becoming a very important field of study and taxonomy is the first and foremost input into this emerging discipline.

Biology at molecular and cellular levels has to be funded, particularly because in India there are some outstanding performers at two or three centres in the country. India has a major stake in this area so as to acquire bargaining power while buying and selling biotechnology and its products. However, the general clan of biotechnologists is not likely to make any worthwhile contribution in the foreseeable future despite heavy investments. This is clear from a critical appraisal of the work done so far with not much on the horizon. However, biology at a grosser level poses equally interesting problems which are likely to be of immediate benefit to science and society. Bawa[8] has rightly concluded that in India world-class work has been done in the past and can be done in future also in such areas involving much less fiscal inputs but with equal if not greater utility.

Partly, the decline in taxonomy is the result of apathy of the biologists themselves, who have willingly taken to the bandwagon of second-rate environmental and biotechnological work. Then followed the increasingly lower priority given to taxonomy and allied subjects not only by the Ministry of Environment and Forests (MoEF) but also by the University Grants Commission, CSIR and ICAR, Department of Science and Technology (DST), etc. According to Father C. Saldanha, taxonomy is unfortunately now regarded as 'tax-on-me'. Such an image of this subject must change at the earliest.

In USA, thanks to the Systematics Agenda 2000 (SA 2000)[9], taxonomy has got a shot-in-the-arm and is poised to become a money-spinner. Proper identification of taxa is of fundamental importance and is going to be a lucrative business on account of tremendous stress on biological products.

SA 2000 is an important study which involved three American Societies of taxonomists and systematists supported by the US National Science Foundation. Initially, they resolved themselves in 27 Standing Committees involving over 300 scientists representing a broad array of institutions and specialities. Their reports were reviewed critically both nationally and internationally. The result was identification of three interrelated missions: namely, (1) to discover, describe and inventorize global species diversity (i.e. discovering biological diversity); (2) to analyse and synthesize the information derived from this global discovery effort into a predictive classification system that reflects the history of life (i.e. understanding biological diversity); and (3) to organize the information derived from this global programme in an efficiently retrievable form that meets the needs of science and society (i.e. managing systematic knowledge).

The value of the study lies in the use of biodiversity for human health, medicine and pharmaceuticals, agriculture, genetic resources, forestry, fisheries, species economics, understanding and conserving earth's life support system, enhancement of quality of everyday life and scientific

research base, etc. In short, this effort is an initiative to discover, describe and classify the biota of the world.

In India, special attention has also to be paid to the marine biodiversity, including mangroves. Let us not forget that our country has been endowed with a long coastline and an extended economic zone (EEZ) which is nearly as large in area as the country itself. We know almost next to nothing about this treasure trove.

It is high time that Indian authorities (especially DST and MoEF) consider bringing together a group of informed biologists to do some deep introspection and chalk out a strategy to halt the decline in taxonomy, and revitalize the subject so as to meet the newer and unforeseen challenges of the future.

Biological Survey of India

Another task that merits attention of such a committee is to bring together BSI and ZSI into one connected whole, as Biological or Ecological Survey of India, where botanists, zoologists, ecologists, conservation biologists and others using the most modern tools and skills, including space science, bring out a consolidated picture about Indian biota, their distribution, composition, endemism, rarity, extent and nature of bioenrichment/biodepletion, early warning signals and 'hot spots' of the Indian biota in space and time. Furthermore, based on their study, they establish trends so that meaningful action can be planned for an overall ecological restoration in general, and restoration of endangered species in particular. Gone is the time when the members of the five kingdoms (Monera, Protista, Fungi, Animalia and Plantae) occupying a common habitat and existing as in interacting system are treated as bits and pieces. One cannot delink a flowering plant from its pollinator or mycorrhiza, a carnivore from its prey, a herbivore from herbage that it eats, above all different biota from their common habitat, and numerous other associations. There is a need to understand processes like mutualism, co-adaptation, co-evolution, predation/parasitism and competition, and speciation together with fundamental processes of mutation, recombination and natural selection leading to survival of the fittest. Equally important are the ecological processes and the role of people living harmoniously in an ecosystem, and especially the lessons to be learnt from such associations.

This does not mean that there is a need to disband the staff of the Surveys, but the need is to revitalize and establish a cadre of taxonomists who look at all biota in a given habitat as an interacting, interconnected, interrelated, interdependent and holistic system rather than as isolated bits and pieces. This would need creating a new brand of dedicated taxonomists who do not look at plants, animals and microorganisms with a dry lens and a dry mind, but as ecologic and economic entities. There will have to be specially designed and designated training centres. This has to be done even if it requires the advice and help of foreign experts and agencies. In time to come, enormous gains would accrue out of this exercise. Advanced genetics, breeding, biotechnology and pharmaceutical sciences have to be backed by advanced taxonomy, both classical and experimental.

1. Ghosh, A. K., *Animal Resources of India: Protozoa to Mammalia, State of the Art*, Zoological Survey of India, Ministry of Environment and Forests, 1991, XI–XXVII.
2. Khoshoo, T. N., *Curr. Sci.*, 1994, **67**, 577–582.
3. Singh, N. P., Conservation of biodiversity, Second Training Course on Management of Natural Resources and Environment, Botanical Survey of India, Ministry of Environment and Forests, Government of India, 1994.
4. World Conservation Monitoring Centre (WCMC), *Global Biodiversity*, Chapman Hall, London, 1993.
5. Khoshoo, T. N., in *Indian Geosphere and Biosphere* (eds Khoshoo, T. N. and Sharma, M.), National Academy of Sciences, Allahabad, 1990, pp. 178–233.
6. Boyden, S., *Biohistory*, The Parthenon Publishing Group, Paris, 1992.
7. Morin, N. R., *et al.*, *Floristics for the 21st Century*, Miss. Bot. Garden, USA, 1988.
8. Bawa, K. S., *Curr. Sci.*, 1993, **64**, 205–207.
9. Systematics Agenda 2000: Charting the Biosphere, New York, 1994.

ACKNOWLEDGEMENTS. This paper represents contribution number 23 of the Research Programme on Conservation of Biodiversity and Environment jointly organized by TERI and the University of Massachusetts at Boston. The programme is supported in part by McArthur Foundation.

T. N. Khoshoo is in the Tata Energy Research Institute, India Habitat Centre, Lodi Road, New Delhi 110 003, India.

SCIENTIFIC CORRESPONDENCE

Record of opal phytoliths and associated palynofossils from Indian crude oils

Opal phytoliths are recorded from the crude oils along with palynofossils. These indicate the Paleogene source for the Bassein crude oils of the Bombay offshore area. The phytoliths, the associated palynofossils and dominant humic organic matter suggest a terrestrial origin for the crude oils.

Five new phytoliths are recorded and described. They are grouped as oval, triangular, ellipsoidal, circular and polygonal phytoliths.

The study of crude oils is an effective tool to determine the age of source rock and source organics of these oils. It is termed as oil palynology. Different types of spores, cellular structures, cuticular structures and wood fragments have been extracted earlier from the crude oils of Digboi field of Assam, India, to find out the source of the oils[1]. Panicoid and elongate type of phytoliths have been recorded from the subsurface sediments of North Bassein Well No. A, Bombay

offshore area, and the historical background of phytolitharians has been discussed[2]. Microfossils have been extracted from the nine crude oils of Jurassic, Triassic and Permian reservoirs in Australia to explain the controversy of Permian/Jurassic origin of oil. Crude oil studies have demonstrated that the oils were derived from the Jurassic source and not from a Permian source[3]. The presence of the spore and pollen assemblages in hydrocarbons from southern Israel indicates a common generation and migration in all the sampled oils[4]. It is concluded that the study of crude oils plays an important role in investigating the age of source rock and the time of migration of the hydrocarbons. The palynological assemblages have also been recorded from the crude oils of the buried hills of China[5]. They discussed the dating of the source bed, oil and gas migration and oil source in the paper.

Kolbe[6] suggests that the better oil sources have better number of fossils in the crude oil and *vice versa.*

The opal phytoliths and other palynofossils were recorded from three crude oils. The crude oils A, B and C were obtained from BA-A, BA-B and BA-C wells, respectively. The wells are located on south Bassein structure of the Bombay offshore area (Figure 1). The crude oils were diluted by using suitable solvents. Then the residual matter was recovered by sieving the solution through a 8–10 μm sieve. The solution was again centrifuged on higher RPM for the recovery of macerate. The solvents used were: xylene (hot/cold), chloroform, petroleum ether and absolute alcohol. The recovery of the macerate was meagre. Permanent slides were prepared by the usual palynological technique and examined under the Leitz microscope using oil immersion objectives.

The main objective of the study was to record phytoliths; however, it also yielded palynofossils, organic matter and other organic remains. The slides are deposited in the repository of Palynology Laboratory, KDMIPE, ONGC Ltd., Dehradun, numbered as IPE/PALY/93/OILA(1-3), OILB(1-4), OILC(1-5).

Phytoliths are the silicified casts of epidermal cells of terrestrial plants, mainly grasses[6,7]. Indian corn produces several morphotypes of panicoid and elongate phytoliths[8]. Subsequently, these microscopic bodies of opaline silica, are seen in seeds, fruits and bracts of many families of angiospermous higher plants of tropical climate. The role of phytoliths is increasing tremendously for providing an independent source of palaeoecological interpretation. They have limited biostratigraphic values[9,10].

The nomenclature of the phytoliths is unsettled; however, five classes have been proposed in the literature[11,12]. They are

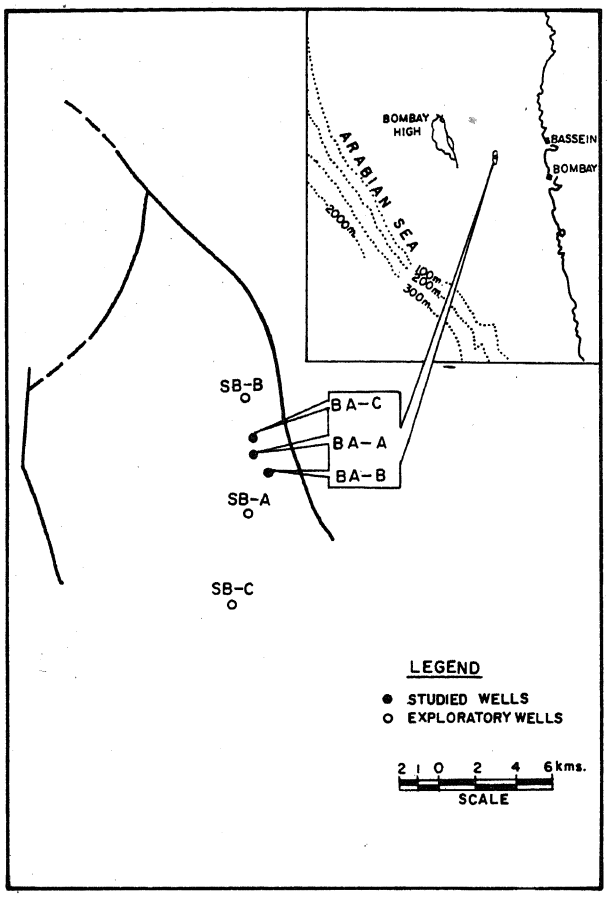


Figure 1. *a*, Index map of Bombay offshore area, India, showing the South Bassein Structure. *b*, South Bassein Structure showing well locations BA-A, BA-B and BA-C, from where crude oils were studied.

Table 1. Distribution of phytoliths and palynofossils in the crude oil A, B and C of Bombay offshore area

Phytoliths and associated palynofossils	Crude oil		
	A	B	C
Panicoid phytolith	+		
Oval phytolith	+		+
Triangular phytolith	+		
Elliposidal phytolith	+	+	+
Circular phytolith	+		+
Polygonal phytolith		+	+
Fungal spore	+	+	
Trilete spore	+	+	+
Proteacidites sp.	+		
Polyvestibulopollenites sp.	+		
Palmaepollenites sp.	+	+	
Humic organic matter	+	+	+
Collenchymatous plant cells	+	+	+
Sclereids with projections	+		
Lignified trichomes	+		

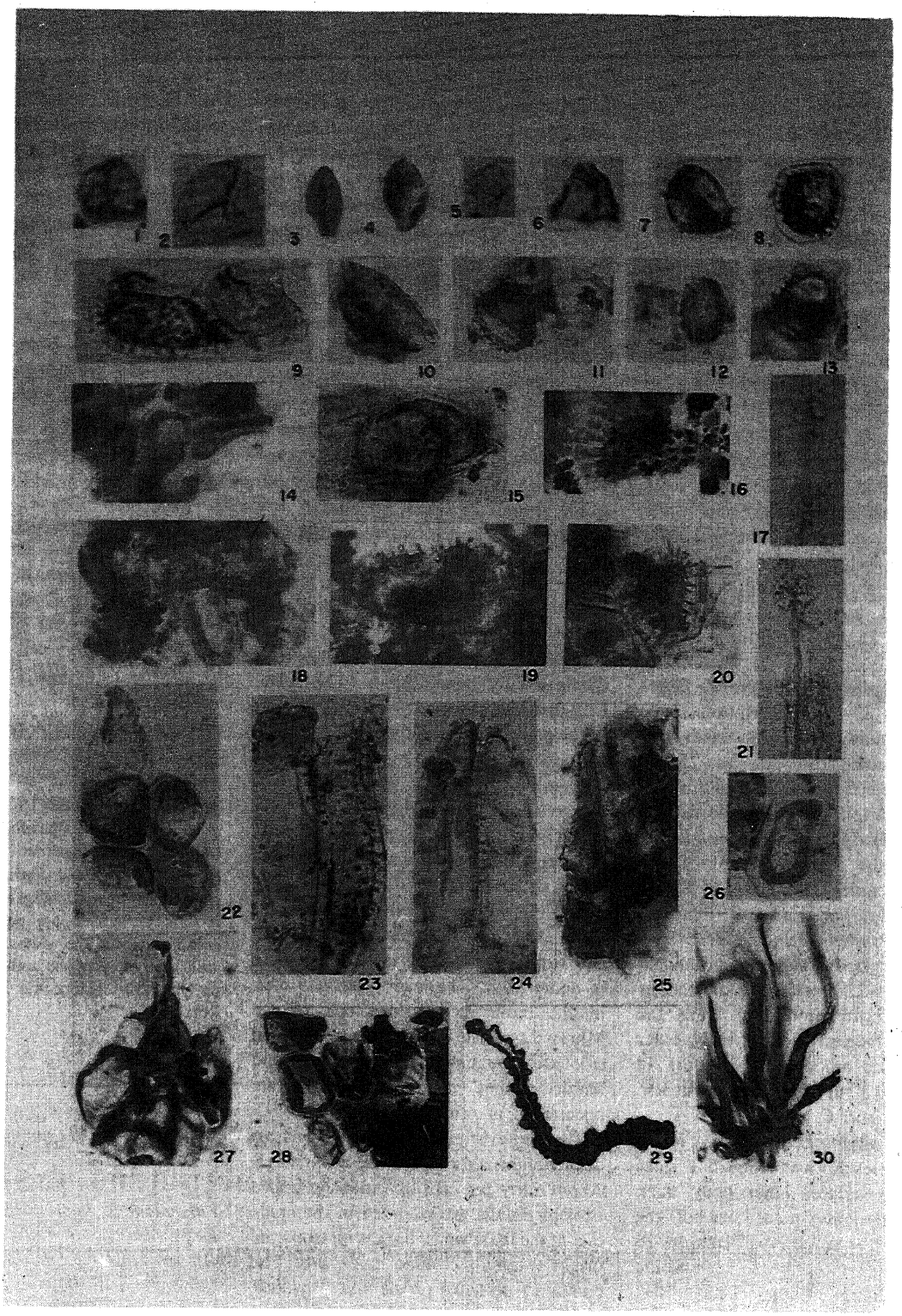


Figure 2. Explanation of photomicrographs. All are × 800 magnified. **1, 2**, trilete spore, crude oil B (1) and A (2); **3**, fungal spore, crude oil A; **4**, *Palmaepollenites* sp., crude oil A; **5**, *Polyvestibulopollenites* sp., crude oil A; **6**, *Proteacidites* sp., crude oil A; **7, 12**, oval phytolith, crude oil A (7) and C (12); **8, 11, 13**, circular phytolith, crude oil A (8) and C (11, 13); **10**, triangular phytolith, crude oil A; **14, 15, 26**, polygonal phytolith, crude oil B (14) and C (15, 26); **19, 20, 23–25**, ellipsoidal phytolith, crude oil A (9), B (16, 19) and C (20, 23–25); **17, 21**, panicoid phytolith, crude oil A; **18, 22, 27, 28**, collenchymatous plant cells, crude oil C (18), A (22, 27) and B (28); **29**, sclereids with projections, crude oil A; **30**, lignified trichomes, crude oil A.

grouped as follows: (1) festucoid, (2) chloridoid, (3) panicoid, (4) elongate, (5) agrostoid and (6) nardoid. Out of these six phytoliths only panicoid-type were observed in the present study. The distribution of phytoliths and the associated palynofossils is shown in Table 1. Five new types of phytoliths and the panicoid phytolith are illustrated in Figure 2. A brief morphological description of these new groups of phytoliths and their occurrence is presented below.

Oval phytolith (Figure 2, Plates 7, 12). Description: shape oval; inner body thick, prominent and having size 12.5 µm; 1.3 µm long spinous projections at the periphery; outer opaline surface 2.5 µm and the overall size 16–22.5 µm. Occurrence: crude oils A and C.

Triangular phytolith (Figure 2, Plate 10). Description: shape triangular; inner body prominent having maximum length 34 µm; 4–5 small projections of about 0.5 µm at the periphery of the inner body; outer opaline surface smooth and 2 µm thick; overall longer side 36 µm and smaller side 25 µm. Occurrence: crude oil A.

Ellipsoidal phytolith (Figure 2, Plates 9, 16, 19, 20, 23–25). Description: shape ellipsoidal; inner body prominent and ranging in size from 20 to 75 µm; numerous projections of 2.5–3.8 µm size seen at the periphery and on the surface of the inner body; outer opaline surface smooth and the overall size 31–80 µm. Occurrence: crude oils A, B and C.

Circular phytolith (Figure 2, Plates 8, 11, 13). Description: shape more or less circular; inner body thick and prominent size 18 µm; 8–22 small-sized projections of about 0.5 µm protruding from the peripheral surface of the inner body; outer opaline surface smooth and ranging in size from 22.5 to 25 µm. Occurrence: crude oils A and C.

Polygonal phytolith (Figure 2, Plates 14, 15, 26). Description: It has 5–6 sides in the outer surface. Inner body thick and prominent; maximum size of the inner body in the range 20–37 µm; 10–16 projections of 1–3.5 µm size seen at the periphery of the inner body of phytoliths; the two sides of the outer surface at junction are projecting; overall size including projections is in the range of 29–45 µm. Occurrence: crude oils B and C.

Palynofossils. Palynofossils are documented from crude oil. The distribution is shown in Table 1 and illustrated in Figure 2. Trilete spore (Plate 1, 2); *Proteacidites* sp. (Plate 6); *Polyvestibulopollenites* sp. (Plate 5); *Palmaepollenites* sp. (Plate 4); fungal spore (Plate 3).

Organic matter. Humic organic matter is present in all the three crude oils. It has not been illustrated.

Other assemblages. Collenchymatous plant cells, common (Figure 2, Plates 18, 22, 27, 28); sclereids with projections, rare (Plate 29) and lignified trichomes, rare (Plate 30). For distribution see Table 1.

The assemblages of phytoliths and palynofossils are diversified in crude oils A and C as compared to crude oil B (Table 1). A common to good percentage (30%) of phytoliths in crude oils indicates probably one genesis for these studied crude oils. The organic matter yield is dominated by humic facies (40%). The frequency and diversity of palynofossil is comparatively low (15%). The presence of palynofossils *Proteacidites* sp., *Polyvestibulopollenites* sp. and *Palmaepollenites* sp. suggests that crude oils are derived from Paleogene source rocks of terrestrial nature, which are mainly shales. It indicates a wide age range. The dating is tentative due to the paucity of data. Other assemblages constitute 5% and organic matter of sapropelic nature is about 10%. The main reservoir rock in the basin is the limestone of the Miocene age. The constituents in the crude oils are dominated by phytoliths and humic organic matter and show rare occurrence of other palynofossils.

The study is based on the occurrence of phytoliths and palynofossil assemblage. About thirty per cent of variously shaped opal phytoliths are recorded in the crude oils of the three wells, indicating terrestrial origin for all the three crude oils. The combined presence of the opal phytolith, humic organic matter, palynofossils and other terrestrial forms shows that the crude oils A, B, and C are terrestrially sourced of Palaeogene age.

Five new groups of phytoliths, viz. oval, triangular, ellipsoidal, circular and pentagonal phytoliths, were recorded from the crude oils of Bombay Offshore Basin, India.

1. Banerjee, D., *Trans. Min. Geol. Met. Inst. India*, 1965, **62**, 51–66.
2. Mathur, K., Proc. X Ind. Coll. Micropal. & Strat., 1982, pp. 501–504.
3. Jersey, N. J. de, *Geol. Surv. Queenstand*, 1965, **329**, 1–9.
4. Horowitz, A. and Langozky, Y., *Geol. en Mijnb*, 1965, **44**, 59–63.
5. Wang, K. F., Zhang, Y. L. and Jiang, H., *Pollen et spores*, 1985, XXVII, 471–489.
6. Kolbe, R. W., *Science*, 1957, **126**, 1053–1056.
7. Floger, D. W., *Deep Sea Res.*, 1970, **17**, 337–352.
8. Twiss, P. C., Suess, E. and Smith, R. M., *Soil Sci. Soc. Am. Proc.*, 1969, **33**, 109–115.
9. Piperno, D. R., *Rev. Palaeobot. Palynol.*, 1985, **45**, 185–228.
10. Piperno, D. R., *Rev. Palaeobot. Palynol.*, 1989, **61**, 147–173.
11. Smithson, F., *J. Soil Sci.*, 1958, **9**, 148–154.
12. Bukry, D., *Init. Rep. Deep Sea Drill. Proj.*, 1979, XLIX, 977–1009.

ACKNOWLEDGEMENTS. We are grateful to Shri Kuldeep Chandra, OSD & Head KDMIPE, Dr Jagadish Pandey, General Manager (GRG) and Shri J. L. Ganju, Dy. General Manager (Geology) for according permission to publish the paper and for their encouragement. The views expressed in this article are those of the authors and not necessarily of the organization in which they are working.

C. M. MISRA
R. K. SAXENA

Palynological Laboratory
KDM Institute of Petroleum Exploration
Oil & Natural Gas Corporation Ltd
Dehradun 248 195, India

Massif-type anorthosite near Ayialur, Dindigul, Anna District, Tamil Nadu, South India

A new anorthosite mass is found as differentiated massif pluton near Ayialur, Dindigul Anna District, Tamil Nadu (longitude between 78°6′ and 78°9′ and latitude between 10°32′ and 10°35′). The massif-type anorthosite mass shows fine-grained border gabbroic facies to coarse-grained anorthosite core through medium-grained gabbroic anorthosite and anorthositic gabbros. The anorthosite mass exhibits foliated fabric in marginal zones with even-grained, granular fabric in core. Ilmeno–magnetite–pryoxenite rocks occur as thin layers in association with anorthosites.

The area is composed of metasediments including metaquartzite, quartz schists, biotite gneiss, amphibolites and calc gneisses. The trends of metasediments reveal two generations of folds with E–W axial trend plunging to the west and NNE–SSW trend plunging to the north.

The anorthosite massif-type body found emplaced as a dome-like mass occupies the antiformal cross-folded structure in the midst of quartzites. It is encircled and caped by metaquartzites and quartz schists. The quartzites trend in ENE–WSW through NE–SW to NW–SE directions, indicating two periods of folding generated in the area.

Petrography

Anorthosites and related rocks are coarse- to fine-grained leucocratic to melanocratic rocks. Anorthosite is chiefly composed of labradorite with subordinate hornblende. Gabbroic anorthosite and fine-grained gabbro contain diopsidic pyroxene. Common accessories are opaques (magnetite + ilmenite) and zircon. There is a general variation in the relative proportion of mafics and felsics from anothosites to gabbros.

Table 1. Estimated modal composition of the anorthosites and related rocks

	Anorthosite	Gabbroic anorthosite	Anorthositic gabbro	Fine grained gabbro
Plagioclase (labradorite)	95.00	78.00	65.00	55.00
Hornblende	3.50	16.50	25.00	30.00
Diopside (relict and discrete phases)	—	3.20	6.50	10.50
Ores (magnetite, ilmenite)	1.50	2.30	3.50	4.50
	100.00	100.00	100.00	100.00

The estimated modal compositions of the anorthosites and related rocks are shown in Table 1. The data show a close agreement with Buddington's classification of anorthositic rocks of Adirondack[1] and Kadavur anorthosites[2].

1. Buddington, A. F., in *Origin of Anorthosite and Related Rocks* (ed. Isachen, Y. W.), NY State Mus. Sci. Serv. Mem., New York, 1968, vol. 18, pp. 215–231.
2. Subramaniam, A. P., *Geol. Mag.*, 1956, **93**, 287–301.

ACKNOWLEDGEMENTS. We express our sincere thanks to Dr R. Mani, Professor and Head, Department of Geology, Annamalai University, for providing facility to do laboratory studies. The second author thanks the UGC for providing the financial assistance.

M. ARUMUGAM
R. SENTHIL KUMAR

Department of Geology
Annamalai University
Annamalainagar 608 002, India

A technique for axenic culture of *Puccinia romagnoliana* – A potential bioherbicide for purple nutsedge

Purple nutsedge (*Cyperus rotundus* L.) is the most troublesome weed in all agricultural production areas of the world having a large vegetative growth capacity through underground tubers. None of the control measures available against this weed are efficient as well as economical. A rust caused by *Puccinia romagnoliana* Maire and Sacc. has been reported for its possible use as biocontrol agent of the nutsedge[1]. Observations with this rust showed significant reduction in tuber number and tuber weight by 34 and 83%, respectively[2]. However, rust being an obligate pathogen, success of developing it into biological control agent depends on many factors, the foremost being the ability to produce abundant inoculum for storage and application in the field. Keeping this in view, an attempt was made to develop axenic culture of *P. romagnoliana*. Different culture media were tested and for the first time a technique has been developed to culture this rust on artificial growth medium.

The leaves of *C. rotundus* were inoculated with rust uredospore suspension,

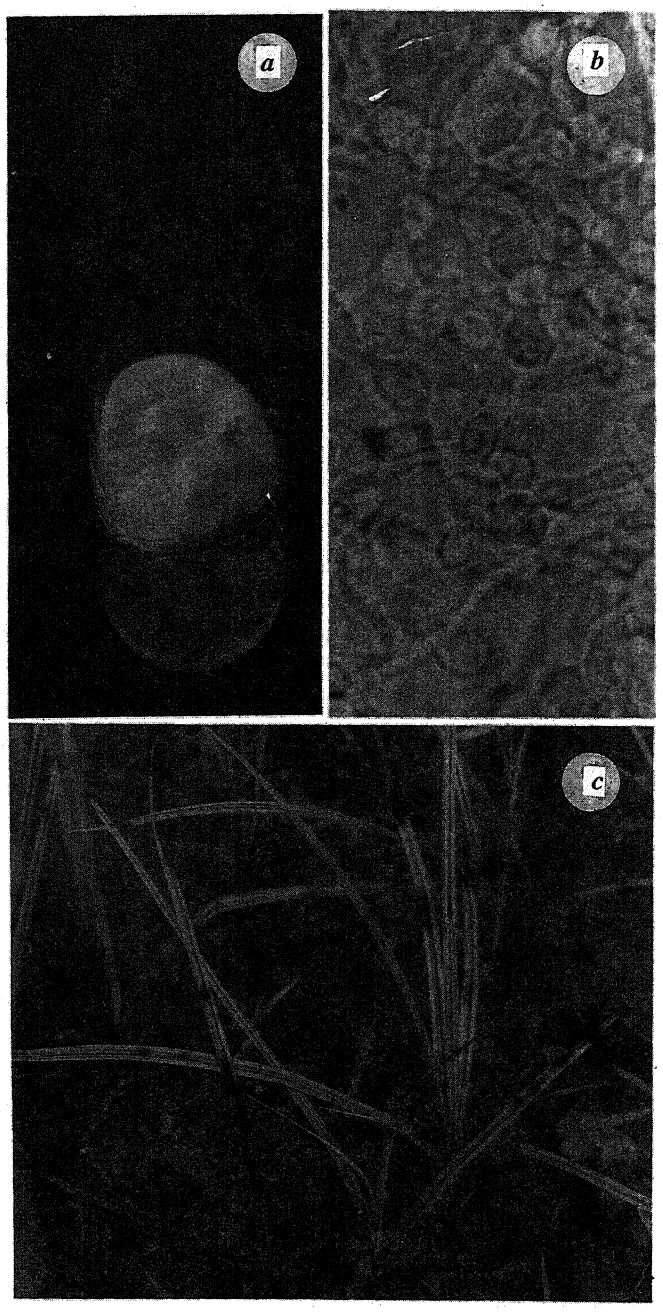


Figure 1 *a–c*. *a*, Axenic culture of *P. romagnoliana* on supplemented MS medium; *b*, Chlamydospore formation of *P. romagnoliana* in culture (225 ×); *c*, Necrotic lesions development on *C. rotundus* leaves after artificial inoculation with chlamydospores of *P. romagnoliana*.

Table 1. Effect of temperature on mycelial growth and chlamydospore formation of *P. romagnoliana*

Temperature (°C)	Initiation of		Colony diameter after 19 days (cm)	Maximum sporulation (days)
	Mycelial growth (days)	Chlamydospore formation (days)		
20	3	4	6.2	22
25	2	3	5.7	14
30	2	4	5.2	18
35	2	4	0.4	9

covered with wet polythene bags for 48 h to provide high relative humidity and incubated at room temperature (approx. 30°C). Olive brown necrotic, nonsporulating lesions were seen to develop in 15–20 days. These lesions were cut into small bits 5–10 mm, surface-sterilized with 0.1% $HgCl_2$ for 2 min, rinsed thrice with sterilized water and plated on MS basal medium[3] (pH 5.8) in petri plates. The medium was supplemented with 2 ppm 2,4-D and 0.5 ppm kinetin. The plates were incubated at 26 ± 1°C. The mycelial growth started in 48 h. A thick stroma of fungus developed on the medium in 5–6 days (Figure 1 *a*). The colony was white in colour, raised and wrinkled and did not penetrate the medium. The cleavage of protoplast and formation of chlamydospores started after 3 days and abundant spores were formed in 10–15 days (Figure 1 *b*). The chlamydospores were initially hyaline, thin-walled but became dark brown and thick-walled in 25–30 days.

Fungus growth and spore formation were studied at different temperatures, viz. 20, 25, 30 and 35°C. Four mm diameter mycelial discs of 10-day-old culture were plated on supplemented MS-basal medium in petri plates and incubated at different temperatures. Maximum mycelial growth was at 20°C (Table 1), whereas at 35°C it was negligible. Chlamydospore formation started after 3 days of incubation at 25°C and after 4 days at other temperatures. Maximum spores, per unit area of mycelial mat, were formed after 14 days at 25°C; although at 35°C it was after 9 days, the mycelial growth forming chlamydospores was very scanty at this temperature.

Plants of *C. rotundus* were artificially inoculated with the chlamydospores obtained in culture. The leaves were rubbed with a wet cotton swab to remove the wax layer, sprayed with chlamydospore suspension, covered with wet polythene bags for 48 h and incubated in isolation under natural conditions (temperature approx. 30°C). Yellow necrotic spots started appearing after 15–20 days which later turned dark brown and necrotic (Figure 1 *c*). These symptoms were similar to the ones from which the fungus was originally isolated and were also observed in naturally infected plants in the field. Reisolations from these spots yielded the same fungus. *P. romagnoliana* significantly reduced the growth of *C. rotundus*.

With the development of axenic culture of *P. romagnoliana* on hormone-supplemented MS medium, one major obstacle in its development as a biocontrol agent against purple nutsedge has been overcome. Although mycelial growth and sporulation are better at 20 and 35°C, respectively, the objective of this study was to obtain maximum sporulation in minimum time. For this, 25°C is the optimum temperature. Further studies are needed to standardize the application technology of this axenic culture.

1. Nitzani, E., Kenneth, R., Kleifield, Y. and Renveni, R., *Phytoparasitica*, 1990, **18**, 240–241.
2. Bedi, J. S., Kaur, Ravinder and Sokhi, S. S., Paper presented at the Global Conference on Advances in Research on Plant Diseases and their Management, 12–17 Feb. 1995, Udaipur.
3. Murashige, T. and Skoog, F., *Physiol. Plant.*, 1962, **15**, 437–497.

J. S. BEDI
RAVINDER KAUR
S. S. SOKHI

Department of Plant Pathology
Punjab Agricultural University
Ludhiana 141 004, India

Polyphenolic components of flowers of *Helichrysum bracteatum*

Helichrysum bracteatum (Asteraceae) is a stout annual herb, naturalized in high altitudes. It is commonly available in the Nilgiris and Kodaikanal hills of south India[1]. The flowers are large, everlasting and appearing in golden yellow, pink and ivory white shades. *Helichrysum* species are rich in secondary metabolites and have good medicinal value[2]. Previous reports reveal that *H. bracteatum* contains lignin and aurones[3,4]. There is no record of isolation of any other constituent. In view of the reported medicinal values of *Helichrysum* species, three differently coloured (golden yellow, pink, ivory white) flowers were examined for their polyphenolics.

Fresh flowers of *H. bracteatum* var. *bracteatum* collected from the Nilgiris hills in south India were refluxed with 90% boiling ethanol and concentrated under reduced pressure. The aqueous crude extract was fractionated into benzene, ether and ethyl acetate solubles. Ether and ethyl acetate soluble fractions tested positive for polyphenolics. These fractions were mixed together, concentrated and kept in an ice chest, where they gave a yellow solid, which on PC and TLC indicated the presence of UV-active compounds. Separation was carried out by preparative cellulose TLC[5] with 50% acetic acid as the developing solvent. The separated zones were eluted with hot methanol and the residue recrystallized from MeOH. The compounds characterized[6] and the methods[7] used are given in Table 1.

^{1}H NMR spectral data (400 MHz, $CDCl_3/CD_3OD$, TMS as int. std., δ, ppm) for the two rare flavonoids showed the following characteristic signals: Bractein: 6.71 (s, 2H, H-2′,6′), 6.31 (s, 1H, =CH–), 6.04 (d, J = 2 Hz, 1H, H-7), 5.98 (d, J = 2 Hz, 1H, H-5), 4.66 (d, J = 8 Hz, 1H, anomeric glucose proton), 3.65–3.17 (m, 6H, other glucose protons); cernuoside: 7.22 (d, J = 2 Hz, 1H, H-2′), 6.98 (dd, J = 8,2 Hz, 1H, H-6′), 6.60 (d, J = 8 Hz, 1H, H-5′), 6.42 (s, 1H, =CH–), 6.09 (s, 1H, H-7), 6.02 (s, 1H, H-5), 4.68 (d, J = 8 Hz, 1H, anomeric glucose proton), 3.68–3.19 (m, 6H, other glucose protons). A survey of the distribution flavonoids in Asteraceae reveals the predominance of apigenin and luteolin along with their 7-glucosides[8]. The present finding of luteolin and its 7-glucoside is in further support. The co-occurrence of phenylpropanoid, aurones and flavonoids in *H. bracteatum* is in agreement with the biogenesis of these types of compounds involving common precursors like mevalonic acid, shikimic acid and phenylalanine. The present study revealed that three morphologically similar varieties of *H. bracteatum*, differing only in the colour of the flower, biosynthesize three different flavonoid compounds – aurone glycosides and flavone in golden yellow, aurone glycosides, and flavone O and C-glycosides in pink and flavone-O-glycoside in ivory white and thus correspond to three chemotypes differentiated by chemical constituents. Bractein was screened for anti-cancer and anti-HIV properties with the help of

Table 1. Compounds isolated and characterized from the flowers of *H. bracteatum*

Compound	Golden yellow	Pink	Ivory white	Characterization method
5,7,3′,4′-tetrahydroxyflavone (luteolin)	+	+	+	m.p., UV, IR, co-spot with authentic sample
6,3′,4′,5′-tetrahydroxy 4-O-β-D-glucopyranosylaurone (bractein)	+	+	–	m.p., UV, IR, ^{1}H NMR, FDMS, hydrolysis
6,3′,4′,-trihydroxy-4-O-β-D-glucopyranosylaurone (cernuoside)	+	+	–	m.p., UV, IR, ^{1}H NMR, FDMS, hydrolysis
(E)-3,4-dihydroxy cinnamic acid (caffeic acid)	+	+	+	m.p, UV, IR, co-spot
6-C-β-D-glucopyranosylluteolin (iso-orientin)	–	+	–	m.p, UV, ^{1}H NMR, co-spot
7-O-β-D-glucopyranosylluteolin (luteolin-7-O-glucoside)	–	–	+	m.p, UV, IR, hydrolysis, co-spot

+, Present; –, Absent; m.p., Melting point.

National Institute of Health, USA. Bractein exhibited a certain amount of activity, but the level was not significant to warrant detailed study.

1. Fyson, P. F., *The Flora of South Indian Hill Stations*, Periodical Expert Book Agency, New Delhi, 1986, vol. 1, pp. 335.
2. Bhakuni, D. S., in *Advances in Medicinal Phytochemistry* (eds Barton, D. and Ollis, W. D.) John Libby Co., Ltd., England, 1986, p. 40.
3. Kisiel, W., *Planta Med.*, 1980, **38**, 285–287.
4. Hansal, R., Langhammer, L. and Albrecht, A. G., *Tetrahedron Lett.*, 1962, 599–602.
5. Jayaprakasam R. and Sivakumar, R., *Curr. Sci.*, 1989, **58**, 980–982.
6. Nair, A. G. R. and Sivakumar, R., *Phytochemistry*, 1990, **29**, 1011–1013.
7. Mabry, T. J., Markham, K. R. and Thomas, M. B., *Systematic Identification of Flavonoids*, Springer-Verlag, New York, 1970.
8. Harborne, J. B. and Mabry, T. J., *The Flavonoids: Advances in Research*, Chapman and Hall, London, 1982.

ACKNOWLEDGEMENTS. Our thanks are due to Dr Ven Narayanan and the National Cancer Institute, Bethesda, Maryland, USA, for the screening of our compound for the anti-cancer and anti-HIV activity and to Mr Mohan, Pandian Estate, Nilgiris, S. India, for the supply of plant material.

R. SIVAKUMAR*
A. G. RAMACHANDRAN NAIR
K. P. MADHUSUDHANAN†

Department of Chemistry
Pondicherry University
Pondicherry 605 014, India
†Department of Medicinal Chemistry
Central Drug Research Institute
Lucknow 226 001, India
**Present address: Salim Ali Centre for Ornithology and Natural History*
Kalampalayam, Coimbatore 641 010, India

MEETINGS/SYMPOSIA/SEMINARS

1995 Society for Industrial and Applied Mathematics (SIAM) Annual Meeting

Place: Adam's Mark Hotel
Charlotte, North Carolina
Date: 23–26 October 1995

Meeting themes include: Computational science and engineering, Education-CSE Degree Programs, Multidisciplinary design optimization, Computational biology, Control of large systems, Numerical algorithms, Numerical software, Environmental modelling.

Contact: Society for Industrial and Applied Mathematics
3600, University City Science Centre
Philadelphia, PA 19104–2688
Phone: 215-382-9800
Fax: 215-386-7999
e.mail: meeting@siam.org

Institution of Chemists (India): Associateship Examination, 1996

The Associateship Examination of the Institution of Chemists (India) will be held in November 1996. The last date for registration is 30 November 1995. For further enquiries regarding this and for Membership:

Contact: The Honorary Secretary
Institution of Chemists (India)
11/4, Dr Biresh Guha Road
Calcutta 700 017

Dr Vishwanath Memorial Fund Award of Gold Medal for Research in Modern Biology

Nominations are invited by the Dr Vishwanath Memorial Foundation for the gold medal award in 1995 for the best research work carried out in the area of modern biology (basic biomedical research) by the scientists working in the Indian institutions, universities, etc. Nominations may be made by Heads of institutions and universities, scientists and individuals accompanied with a biodata and a list of publications be sent to the following address before 31 July 1995.

Dr P. D. Gupta
Centre for Cellular & Molecular Biology
Uppal Road, Hyderabad 500 007
Fax: 040- 671195
e.mail: pdgupta@ccmb.uunet.in

Formation of the International Society for Quality of life

Newly formed professional organization proposes to promote the exchange of information about quality of life and its investigation throughout the world. The Society offers an exciting new opportunity for collaboration, exchange of ideas, collegiality, etc. Membership is available to any individual who shares objectives of the Society. For further information

Contact: Riverwood Association office
Phone: 206. 989. 9025
Fax: 206. 616. 3135
e.mail: darian@uu.Washington.edu

Reviving experimental physics in the universities – The IUC initiative

S. Ramaseshan

*There is widespread concern about the state and future of science in the universities – witness the recent presentation and discussion of the National Science University (NSU) proposal in these pages. One of the measures taken by the University Grants Commission (UGC) to provide improved opportunities for research level work in the university sector is the creation of several Inter-University Centres where unique facilities are shared by workers from all over the country. The Inter-University Centre for Astronomy and Astrophysics (IUCAA) is an example and a brief report by J. V. Narlikar appeared in Current Science some time ago (1993, **65**, 730). The author of this article recently visited the Inter-University Centre for DAE facilities (IUC-DAEF) located on the university campus in Indore and was struck, as many others before him have been, by the many unique features of this institution, where a serious effort is being made to revive first-rate experimental condensed-matter physics in the universities, in a quiet, low-key, low-publicity style. The following sections on the origin of this institution and its facilities and activities have been compiled based on detailed information kindly provided by the staff at IUC-DAEF.*

Origin of IUC-DAEF

It is undeniable that experimental research in physics in universities in India, barring a few notable exceptions, is in very dire straits today. Condensed-matter physics is an area in which experimental work of reasonable quality can be pursued by small teams with not too expensive equipment. Though the majority of the existing universities do pursue research in condensed-matter physics, facilities for characterization of the materials prepared and for carrying out a reasonably complete range of experiments on materials at different temperatures are very often not available to most researchers. This means that the standard of the research as well as the stimulation that it offers to students and faculty at the university both suffer.

On the other hand, first-rate experimental facilities are often established outside the university system. For example, DAE has built a variety of neutron spectrometers on the Dhruva reactor. Protons, deuterons and alpha particles can be accelerated to several MeV energies with the Variable-Energy Cyclotron in Calcutta (VECC). There are plans to build a superconducting cyclotron. The Centre for Advanced Technology (CAT) in Indore is developing synchrotron radiation sources Indus 1 and Indus 2. It was felt both in the DAE and outside that these frontline research facilities must be available to university scientists. University scientists must be slowly inducted to work with scientists from DAE in developing different types of accelerators, beamlines, etc. One can add that elsewhere in the world this kind of university participation is a rule rather than an exception. The Very Large Array Telescope, which in many senses the world's most powerful radio telescope, is even run by an association of universities, and many accelerators like CERN have large teams of university researchers building and using giant detectors.

Aims

To improve the quality of experimental research in physics in universities and to provide access to the major DAE facilities to the researchers from universities, the UGC set up the Inter-University Consortium for DAE Facilities (IUC-DAEF) in 1990. This is an autonomous society fully funded by UGC. IUC-DAEF has three centres, at Bombay, Calcutta and Indore, the last one being the headquarters.

The IUC-DAEF supports projects from university researchers to use the major national DAE facilities at BARC and VECC and also the upcoming DAE facilities at CAT. To make the best use of these expensive facilities, it is necessary to work with well-prepared and characterized materials and to carry out other experiments which will be complementary to those obtained with the major DAE facilities such as the neutron

S. Ramaseshan is in the Raman Research Institute, Bangalore 560 080, India.

spectrometers at BARC or cyclotron at VECC. Therefore, the IUC-DAEF has the task of setting up supporting facilities. The IUC-DAEF also participates in certain development projects of DAE. Periodic workshops on current topics of research for the benefit of young faculty and research scholars of universities are another important activity.

Utilization of DAE facilities

In the last four years the number of university projects on the Dhruva reactor at BARC has increased to 30. There is a steady stream of university scientists visiting BARC for carrying out experiments with the neutron spectrometers. The majority of the projects use the profile analysis powder diffractometer, the small-angle neutron scattering spectrometer, the high-Q diffractometer for the study of amorphous materials and the neutron spectrometers for magnetic studies, in that order. The SANS (small-angle neutron scattering) spectrometer is being used by chemists from different universities. The IUC-Bombay centre conducts each year in association with BARC scientists a three-week workshop in which about 20–25 scientists from universities and other institutions participate. Two weeks are devoted to lectures on various aspects of neutron scattering and diffraction. In each workshop one aspect of neutron scattering is emphasized. During the discussions that follow the lectures and tutorials, the participants are encouraged to come out with ideas for projects. In the third week of the workshop the participants are taken to BARC so that they can see actual demonstrations of the different spectrometers. This method of generating proposals has proved effective. The proposals are often prepared after a detailed discussion between the investigator from the university and the scientists from BARC. Every year the progress in a project is reviewed at a Project review meeting at which all project investigators and scientists from the Solid-State Division of BARC are present. From 1992–1994 ten papers have been published in journals on results obtained in these projects. The number of presentations at the annual DAE symposium on solid-state physics is 20. Till recently K. R. Rao and P. R. Vijayaraghavan, with the cooperation of their colleagues in the Solid-State Physics Division of BARC, were looking after the administrative and scientific aspects of these projects. Dhananjai Pandey joined as the Centre Director of the IUC-Bombay Centre in the middle of 1994.

In Calcutta, activities started from February 1992, when S. N. Chintalpudi was appointed the Centre Director. There are now 32 projects on the Variable-Energy Cyclotron, including projects in nuclear physics, materials science, chemistry and biology. These projects are also reviewed each year. Seven publications have appeared in international journals and eleven papers have been presented at national symposia so far. Discussion meetings on pre-equilibrium phenomena and in-beam γ-ray spectroscopy have been held to evolve group projects.

Starting from 1993, seven projects were sanctioned on the low-energy accelerators at IGCAR, Kalpakkam. This was done after holding a workshop on the type of experiments that can be done with low-energy accelerators. The projects were reviewed in November 1994. Students from different universities have stayed for extended periods in Kalpakkam and have helped in setting up the various beamlines on the accelerators, and it is expected that useful scientific results will soon be forthcoming.

IUC Indore

The Indore Centre started in July 1990 with V. G. Bhide in charge and R. Srinivasan took over the Directorship in December 1990. This Centre, which is also the headquarters, has independent materials preparation and characterization facilities as well as supporting facilities of EXAFS, ESCA, soft-X-ray spectroscopy, transport properties and AC susceptibility measurements, etc. The centre is involved in developing beamlines for synchrotron radiation for users from universities. One current project is to build a beamline for photoelectron spectroscopy (PES) on Indus 1. The wavelength range in which this PES beamline will work is from 60 to 1600 Å. After an evaluation of the various types of gratings, mountings, etc., it was decided to use a toroidal grating monochromator with three interchangeable gratings to cover this wavelength range. To get adequate resolution for PES work the TGM 2600 design of Bessy has been chosen. The toroidal grating monochromator has been purchased from Jobin Yvon in France and the turbo and sputter ion pumps were attached to it, and ultra high vacuum achieved. The grating has been aligned and the expected resolution checked with He I and He II sources. The design for the mirror chambers, the laser alignment chambers, etc., have been made by scientists from IUC. These are under fabrication. It is expected that the entire beamline, which will be 12.1 m in length, will be assembled and tested by December 1995 in conformity with the expected time schedule of completion of Indus 1. To carry out some basic photoelectron spectroscopy experiments the IUC-DAEF has given a project to D. D. Sarma and M. S. Hegde of IISc, Bangalore, to construct a set-up consisting of a sample preparation chamber, sample analysis chamber, and analyser with provision for putting a twin X-ray anode and a He UV source. This spectrometer has been constructed and is being assembled.

With this beamline one can study the core level

spectroscopy of K levels of light elements, L levels of elements of medium atomic weight and M levels of heavy elements. A discussion meeting on the type of projects to be undertaken has been held in Indore.

The second major project is to build a complete soft-X-ray system to study transmission and reflection of thin films at different angles of incidence. This is of interest from the point of view of mirror coatings for the synchrotron radiation source. This project is partially supported by DST. In this project the rotating magnetic seal with an aluminium anode has been imported from Rigaku and the filament assembly, the vacuum chamber and the power supply have been built locally. It is designed to operate at 3 kW. The soft-X-ray continuous radiation coming out of the rotating anode source will be monochromatized with a double crystal monochromator with a fixed entrance and exit slit. Such a monochromator operated by a single screw driven by a DC motor has been designed and fabricated with a positioning accuracy of 5 μm. The large d spacing crystals have been imported. The channeltron detector has been imported and successfully tested. The experimental chamber with $\theta - 2\theta$ mount is currently being fabricated. The whole system will be operational by October 1995.

Materials preparation and characterization

The IUC-Indore centre has set up materials preparation, characterization and measurement facilities listed below:

1. An EXAFS (extended X-ray absorption fine structure) spectrometer with the mechanical movements designed in Pune University and built by a company there. This is a bent-crystal Johansson-type spectrometer using a Si crystal. This spectrometer is mounted on one window of a 12 kW rotating anode X-ray generator. Since the generator has Cu and Co anodes, EXAFS for elements whose absorption edges fall in the appropriate wavelength range can be carried out. This instrument has been used to study the EXAFS of copper-containing organic complexes, copper-containing chalcogenide glasses, As–Se glasses and glassy bismuth copper oxide materials. Users are from universities, IITs and the IISc.

2. A powder diffractometer mounted on the second window of the Rigaku rotating-anode generator. This is equipped for use at temperatures varying from 80 to 1000 K and allows grain size determination, strain determination, etc. This has been widely used by several university groups to characterize their samples. Last year nearly 500 diffractograms at different temperatures were taken for university users.

3. An ESCA unit from VSW, UK, has been functioning for the last three years. This has provision for XPS with a twin-anode source, UPS (ultraviolet photoemission spectroscopy) with He I and II light source and an electron gun for Auger spectroscopy. There were several problems with this instrument, but all the problems have been solved with the expertise available in Indore. This facility has been availed of by nearly 15 university users so far, mainly for core-level spectroscopy work. It is planned to improve the facility by providing a scraper unit, a variable-temperature attachment and a high-energy Argon ion gun for depth profiling.

Facilities developed in-house

4. A fully automated AC susceptibility set-up with variable temperature in the range 80–300 K and with provision for recording the in-phase and 90° out-of-phase components up to the seventh harmonic. With this, studies have been made on the AC susceptibility of sintered pellets of high-T_c materials. The set-up can be used to study magnetic ordering and the growth of harmonics close to the ordering temperature.

5. An automated low-frequency AC conductivity set-up with a precision of 1 part in 10^5 with temperature control to 0.1 K in the range 80–300 K.

6. A dielectric constant and loss set-up to work in the temperature range 80–600 K and in the frequency range 10–13 MHz. This is being automated.

7. A low-temperature resistivity set-up built on a closed-circuit refrigerator. The resistivity of three (metallic or semiconducting) samples can be measured at the same time with temperature controlled to 10 mK in the range 15–300 K.

8. A low-temperature thermopower set-up mounted on a closed-circuit refrigerator and automated to operate between 15 and 300 K. The thermo-emf can be measured to a precision of 0.1 μV. A set-up for high-temperature thermo-emf measurement is available. A provision for evacuating the sample chamber and filling it with an inert gas needs to be made.

9. A vibrating-reed set-up for measuring internal friction and elastic constant is working and has been automated.

10. An adiabatic pulse calorimeter has been set up and automated to work in the range 80–300 K. This requires samples of 0.5 g to set a good accuracy.

Very soon the temperature range in which the above facilities will be operational will be extended to 4.2 K with cryostats fabricated by a Madras firm.

In addition, there is a ball mill, temperature-controlled resistive furnaces, and an arc furnace for preparation of samples. A thin-film coating unit is also available.

Nearly 60 groups from various universities, IITs and IISc have used the facilities in the last two years.

Low-temperature facility

The entire low-temperature facility for production of cryogens has been set up using second-hand machines obtained as gift from IIT, Madras, the Free University, Berlin, and the Max-Planck Institute for Solid-State Research in Stuttgart. A 25-year-old nitrogen plant obtained as a gift from IIT, Madras, has been installed and has produced in the last two years 8000 l of LN_2. An old helium plant obtained from Free University, Berlin, was commissioned in May 1994 and is working satisfactorily. The IUC has supplied liquid helium to NCL, Pune, twice (60 + 70 l).

In Calcutta facilities have been set up for making targets for users of the cyclotron and many university groups have used these facilities. Temperature-controlled diffusion furnaces for making Si(Li) detectors have been set up and some detectors made. Their performance is being evaluated. Several NIM electronic modules are made available for users of the VECC. A Mössbauer spectrometer has been commissioned and is being regularly used by scientists from different institutions around Calcutta. A positron annihilation set-up and a time differential perturbed angular correlation set-up are being established with the help of several university groups. Laboratories for radiochemistry, general chemistry and toxic chemistry have been set up. Attempts to prepare Mössbauer and positron sources are underway.

The three centres together have conducted nearly 30 workshops/discussion meetings/national symposia in the last four years. More than 700 research scholars and scientists from universities and other research institutions drawn from most of the states of the country have participated in these workshops.

Some special features

Even the rather bald factual account of IUC-DAEF activities given above is impressive, given the short time that it has been in existence. But visitors to the Indore Centre quickly learnt that this is only a part of the story. The way in which the centre was built up and is being run has some special features which may be of wider interest. One important aspect was that even before the main building was complete, the first laboratory activities had started. The recruitment was carried out in a phased manner, so that everyone who joined had plenty of work to do. This situation prevails even at present, and is in healthy contrast to others, alas not uncommon in our country, when a large building and recruitment programme precedes active research. It also appears that the administration is small in size, uses modern means of communication, and is charged with helping the scientists procure what they need with the minimum possible delay. The scientists themselves can devote up to 30% of the time to their own research, which is of course a good incentive to keep the equipment in good condition. This is a lesson for those who would set up pure 'facilities' and appoint people to do only routine work. At the same time, the opposite situation of the facilities being all used in-house, with only lip service to outside users, can also arise! There seems to be no risk of this at IUC at present – if anything, it was clear that enormous effort was being put in to see that the university users went away with the data they wanted. Another important point is the presence of IUC's own research students, who, like the staff, are enthusiastic about their work and spend long hours in the laboratories. Many years ago, C. V. Raman described some of our laboratories as 'graveyards of scientific equipment'. The sight of so much equipment well maintained *and fully utilized* would have warmed his heart. It should be mentioned that there is a strong emphasis on the workshop and on fabrication of many things in-house. *The entire low-temperature facility is based on equipment which had been discarded elsewhere.* By making it work, the staff at IUC have earned the right to ask for new equipment in a manner which is again not common elsewhere!

What of the future?

While the achievements of IUC are laudable, one should also utter a word of caution. By providing such ready-made facilities, one may be depriving the researcher from learning to do good experiments himself. In such a case the IUC will not be promoting the growth of expertise in experimental physics. Perhaps the IUC should think of organizing experimental courses for two to three weeks whereby young scientists are made to handle equipment and do their experiments themselves. This may not be easy because it would mean setting up a duplicate set of less expensive facilities for handling by the students and may cost money. But if the culture of doing good experiments is to grow, the UGC *must* find money for such a programme.

The IUC with its expertise must try to involve young students in designing equipment and, if such designs are successful, it must find a way of involving small-scale industry in fabricating and selling such equipment. Unless we try to make our own equipment, Indian science will be captive.

The workshops organized by IUC have generally been well received. IUC pays for the travel and boarding of the participants. Possibly due to financial constraints, too much is attempted to be covered in too short a time and enough reading material is not given to the participants. Also, due to lack of guest house facilities it has not been possible to put the participants and speakers together. This can hinder the development of interactive participation. A guest-house with a sufficient

number of rooms seems to be an essential requirement. If IUC could restrict the topics to be covered and make the lectures truly pedagogical, the purpose of organizing such workshops will be served better.

IUC should put some effort in improving the laboratory curriculum of B Sc and M Sc classes. It is involved in one such education project which is restricted to five colleges of Madhya Pradesh. The project has just started. It will be interesting to watch how this project succeeds.

Finally, it may not be inappropriate to caution IUC about the dangers of doing too many things in too short a time. IUC must now consolidate the gains it has achieved so far. The experiment of getting university people to work together is alien to our prevailing culture. Hence, one must wait for sufficient time for the idea to take root. Sometimes pushing an idea too fast may generate an adverse reaction!

The success of any venture depends on the initiative and dedication of the leadership. IUC has been lucky in that the first two directors have shown qualities of dedication and purposeful action. To nurture IUC, which is a young plant today, into a spreading tree tomorrow with well-entrenched roots, great care must be taken in selecting the next Director of IUC-DAEF. He must be an experimental physicist of repute and must command the respect of the physics community as a whole. He must be deeply committed to the cause of providing service to the university community and to strengthening the bonds between the scientists of DAE and the universities. He must have administrative competence and should keep in check the universal, ever-present tendency for the growth of red-tapism in any organization. It is to be hoped that the IUC-DAEF will be fortunate in having a succession of such leaders.

Modern science and technology: The public perception

V. Govindarajulu

Positive public perception and attitude towards modern science and technology (S&T) is a vital link in the all-round economic, industrial and cultural development of a nation. Public hostility and/or indifference towards modern S&T, on the other hand, is a drag on the growth of economic productivity of any nation. Public policy options, as derivatives of public perceptions, can be the basis for formulating effective and productive development programmes. The results of this DELPHI study, conducted in the State of Kerala, highlights the public perceptions and attitudes towards modern S&T, productivity and industrial development. The public in the State of Kerala possesses a positive and desirable perception that modern S&T is the key to productivity and national development. The results are comparable to the studies done elsewhere and, therefore, universally valid. This study provides a broad spectrum of philosophical and sociological foundations of modern S&T and industrial development. In-depth areawise, sectorwise and disciplinewise impact assessment studies are required before initiating large-scale investment-oriented S&T projects in the states and in the country.

SCIENCE and technology (S&T), nowadays, plays a vital role in the daily routines of our lives. Industrial development, economic productivity and thus national prosperity depend much on the vital role played by S&T. Efficiency and productivity of our nation's economy are functions of social technology of the people. Social technology is the level of knowledge base, entrepreneurship and social consciousness acquired by the people, that continuously helps them to raise the quality output of goods and services per given quantum of resources as input. The gross domestic product (GDP) per capita or per capita income at any given time is the index of the level of social technology of the people and their economic productivity attained. Thus, per capita income of some Asian countries – in 1993, for example, Japan

V. Govindarajulu is in the Regional Research Laboratory, Industrial Estate P.O., Trivandrum 695 019, India.

$ 23,800, Singapore $ 10,450, Taiwan $ 8500, Malaysia $ 2160, China $ 650, India $ 340 – indicates their level of economic productivity, social technology, entrepreneurship, social consciousness and public perception and attitude towards modern S&T. Accordingly, public hostility and/or indifference to S&T is an insurmountable hurdle to nation's economic and industrial development. The economics of underdevelopment and the politics of communalism, which sustain each other, are the characteristics of stagnant and/or low productive economy. And when they are reinforced by the hostile and/or indifferent public perception and attitude towards modern S&T, they lead the country to a negative entropic decline as seen nowadays. On the other hand, the economics of development and the politics of democratic secularism, which are complementary to each other, are the characteristics of increasing orders of productive economy. And when reinforced by a positive public perception and attitude towards modern S&T, they lead the country to a positive entropic growth. One could infer such an all-round growth in the early stages of many civilizations. Therefore, a positive public perception and attitude towards modern S&T is imperative for shaping more effective public policy options[1]. Such effective S&T and public policy options could make powerful guidelines for action plans and accelerated growth of national economy.

Scope and objective

We have initiated this research project to assess and map out public perceptions and attitudes towards modern S&T by DELPHI study in Kerala and help the development agencies in the state(s) and in the country to harmonize various group interests by a process of reconciliation, while implementing large-scale investment-oriented S&T projects.

Methodology

The methods adopted here to assess and map out the public perception are: DELPHI study, i.e. seeking intuitive views of the concerned public, statistical aggregate analysis of the data received and cross-comparison and correlation with similar studies. A structured questionnaire was prepared and circulated to as many as 3000 volunteers in Kerala drawn at random and a 10% response was received, which is considered to be fairly good.

Though many participants were from in and around Trivandrum district, they were the people representing all parts of the state and the country with wide exposure to world developments. Hence, the study results could represent the public perception of the state as well as of the nation. The presentation, made here in two parts, follows a pattern of stating general principles, followed by public perception and attitude as assessed, and then is concluded by derived public policy options as guidelines for future action plans. Part 1 deals with public perception on modern S&T and economy while Part 2 deals with people's perception and attitude towards development issues and action plans.

The respondents

The public who responded to this DELPHI study were the volunteers with tremendous enthusiasm in the social function of S&T. About 88% are males and the rest females. The participants to this study were aged between 20 and 60 years, of which 38% belonged to the age group 40–49 years, followed by 26% in the age group of 30–39 years, 20% in the age group of 20–29 years, 14% in the age group of 50–59 years and 2% in the age group of 60 years and above. Majority of the respondents (56%) were postgraduates, followed by 28% doctoral, 10% graduates and 6% undergraduates. Currently, 44% of the participants are engaged in scientific research, followed by 18% in engineering, 12% in government bureaucracy, 10% in teaching, 4% in management and 12% in other professions.

Part 1: Modern S&T and economy

Given the freedom: The choice of profession

Given the freedom, the respondents would like to accord first ranking (R1) to scientific research as choice of their career and profession, followed by teaching (R2), engineering (R3), social services (R3), journalism and mass communication (R4), business and industry (R5), so on and so forth (Table 1). It indicates the occupation prestige attributed by the public to different professions and careers. Similar trends to occupation prestige, in recent studies, were also attributed by the public in China and United States[2,3]. It indicates the universal nature of occupational prestige attributed by the public.

Table 1. Given the freedom: choice of profession

First five professions	Respondents (%)	Rank accorded
Scientific research	36	1
Teaching	35	2
Engineering	12	3
Social service	12	3
Journalism and mass communication	13	4
Business and industry	25	5

World problems: Role of S&T

World develops by a process of constant struggle, unity and struggle. Constant struggle and unity are the processes of resolving contradictions that exist in nature, society and human thought, thereby raising economic productivity by many orders. These contradictions are perceived as societal problems. Recognition of such societal contradictions itself is half the solution. Therefore, it is heartening to see here that 88% of respondents to this study do recognize that the people of the world in general face innumerable problems. And 86% of them perceive that such problems do pose potential threats to ecology, environment and human habitat. An equal proportion of respondents assert that such threatening problems could provide ample opportunities to S&T and other professions in searching for feasible solutions.

When asked to rank the groups which are capable of providing more feasible solutions to such threatening problems, the majority accorded R1 to scientists (Table 2). It is followed by technologists (R2), engineers (R3), decision makers in business and industry (R4) and decision makers in the government (R5).

The ordering of ranking by the public is precise, as scientists, as a group capable of understanding the basic laws that govern the development of nature, society and human thought, technologists, as a group capable of developing viable and feasible models and systems to advance the knowledge base to the application levels, engineers, as a group capable of converting these models and systems of knowledge base on a social scale; decision makers in business and industry, as a group of social entrepreneurs converting such a knowledge base into productive investment projects, and, finally, the decision makers in the governments to put the politics in the right perspective (i.e. the art and science of making desirable things possible) to mobilize the resources of states, nation and world to convert all such societal contradictions into socially useful productive opportunities. But, as the saying goes, for every economic problem there is a political solution. That means putting politics in command but on the firm understanding of the knowledge base. Therefore, what is needed here is a mix of all disciplines commencing from all directions, each one taking lead from strengths of its own discipline and merging with strong points of others at appropriate stages.

Table 2. Groups capable of providing more feasible solutions

Groups capable of providing feasible solutions	Respondents (%)	Rank accorded
Scientists	35	1
Technologists	29	2
Engineers	34	3
Decision makers in business and industry	31	4
Decision makers in the government	21	5

S&T development: Decision by whom?

The public has given good regard to scientists, technologists and engineers. As these groups are capable of offering more feasible solutions to the world-threatening problems, the public desires scientists, technologists and engineers to take the first three places of responsibility of decision making in the process of S&T development (Table 3). It is also important to see that the bureaucrats in the governments, people in non-governmental organizations (NGOs) and industries are accorded R4th, R5th and R6th, respectively, to hold responsibilities of decision making in S&T development.

But a synthesis of roles of all the groups is necessary to obtain synergestic effects by striking harmony of interests.

S&T: People's motives

People's motives towards certain disciplines are conditioned by certain internal urges inherent to them besides the opportunities for material and spiritual gains. Accordingly, the response here is more positive and corresponds to Maslow's desirable hierarchical theory of motivation. About 48% of the respondents accord R1 to the disciplines of S&T as they give them opportunities for achieving self-accomplishment (Table 4). It is followed by social recognition (R2), excitement (R3) and security and basic needs (R4). It is a surprising but desirable response, in the present days of crisis-ridden and poverty-dominant society, where the majority of the people are generally driven by the basic and security needs. It indicates people's willingness towards self-accomplishment by advancing the knowledge base despite all-round

Table 3. Groups to hold responsibility of decision making in S&T development

Groups	Respondents (%)	Ranking
Scientists	54	1
Technologists	48	2
Engineers and professionals	44	3
Bureaucrats in the government	38	4
Others (NGOs)	16	5
Business and industrialists	14	6

Note: Multiple responses accepted.

Table 4. S&T and people's motives

Motive factors	Respondents (%)	Rank accorded
Self-accomplishment	48	1
Social recognition	38	2
Excitement	27	3
Basic needs and security	35	4

backwardness and limitations. People understand that knowledge is the source of wealth and one should advance it by all means and sacrifices. Therefore, one has to consolidate and strengthen this urge for self-accomplishment by superimposing Herzberg's two-factor hygiene theory of motivation on Maslow's desirable hierarchical theory of motivation. That is achieved by providing a conducive work environment. And such an environment will raise the economic productivity of people at par with that of advanced countries.

S&T progress: Does it benefit modern life?

Four decades back a vast majority of people in the country, particularly in the rural areas, were ignorant, suspicious and even hostile to modern S&T. One could recall such suspicions and hostilities, by the rural folk, towards vaccination to small pox. But today the scene is the reverse. Growth in literacy and awareness through mass media and cultural exchanges by travel and tourism have made all such desirable changes possible.

Accordingly, an overwhelming majority (98%) of the respondents here agree that S&T progress has immensely benefited the mankind and modern life. However, 56% of the respondents assert that such benefits extended to the whole cross-sections of the society, while 44% assert that it covered certain pockets only. Over 54% of the participants to this study assert that S&T progress has brought more of good impacts on society while 40% assert that there are both good and harmful impacts in equal proportions and 6% express nil opinion. Fortunately, none of the respondents attribute the occurrence of harmful societal impacts to S&T progress alone. Hence, the need for consolidating good impacts by weeding out bad ones by proper technology assessment and technology forecasting exercises and mass awareness campaigns. It should also be ensured that the good impacts spread to the whole spectrum of society benefiting one and all.

Societal benefits: The direct derivatives of S&T efforts

People's perception and attitude towards modern S&T is conditioned by their direct experiences with the modern amenities. Thus, public perception of societal benefits, as direct results of S&T efforts, is more realistic and positive (Table 5). About 70–80% respondents attribute to modern medical aids, better transportation and communication network, electrical domestic appliances and family planning and child birth control devices as societal benefits derived directly from the results of S&T efforts. About 50–60% respondents assert that better educational facilities, urban public utilities, declined infant mortality and better food preservation methods are other societal benefits. And 40–50% view that extended life expectancy and modern methods of recreation are direct societal benefits due to S&T progress. This is an index of people's priorities to life-styles and systems, which deserve more attention. Raise in the people's per capita income can enhance the demand for such modern amenities much beyond the present level. Accordingly, cost-effective and more productive technologies are needed to produce such quality amenities at competitive prices.

Societal harmful effects: Inherent or by-product of S&T?

Societal harmful effects, in some cases, are both inherent and by-product of S&T. The first-generation technologies, which were energy-inefficient, exhibited both the characteristics. They get intensified when disregard is given to public safety under the presumption that they do not add to the benefits. But they can be overcome by the present-generation technologies when a balance is struck between social costs and social benefits. Thus, the public expresses its concern and alarm over the present-day harmful societal effects derived as direct results of S&T efforts (Table 6). About 60–80% of the respondents assert that arms race, militarization of industrial economy and environmental pollution and ecological imbalances are the societal harmful effects resulting directly from S&T progress. Another 40–60% respondents attribute the abuse of pesticides, insecticides and chemical fertilizers and excessive dependence on medicines and abuse of drugs as

Table 5. Social benefits as direct results of S&T efforts

Societal benefits as direct results of S&T	Respondents (%) attributing
Modern medical aids	82
Better transport and communication network	78
Electrical domestic appliances	76
Family planning and child birth control devices and systems	66
Better educational facilities	62
Urban public utilities	54
Declined infant mortality	54
Better and effective food preparation and food preservation methods and systems	52
Extended life expectancy	48
Modern methods of recreation	40

Note: Multiple responses accepted.

Table 6. Societal harmful effects as direct results of S&T efforts

Societal harmful effects as direct results of S&T	Respondents (%) attributing
Arms race and militarization of industrial economy	77
Environmental pollution and ecological imbalances	66
Abuse of pesticides/insecticides/ chemical fertilizers	58
Excessive dependence on medicine and abuse of drugs	44
Food adulteration and low-quality foods	34
Degradation and withering of traditional values	26

being direct results of S&T. And 25–40% respondents believe that food adulteration, low-quality foods and degeneration and withering of traditional values are the direct results of S&T progress. Direction of S&T efforts determines its results, whose impacts can be either good or harmful. Sometimes the nature of impact can be the inherent characteristics of the type of technology itself. So what is needed is to do technology assessment and technology forcasting exercises, anticipate the nature of impacts, overcome the negative and consolidate the positives and educate the masses.

Societal problems: The root causes

Societal harmful impacts due to S&T obviously cause societal problems as well. They need to be thoroughly understood and properly tackled. Public perception here is that 38% respondents assert a few societal problems, while 34% think that some societal problems are caused by S&T progress, 6% assert most of the problems, 2% assert no problem and 6% express nil opinion. When asked to identify some such societal problems caused by S&T progress, 58% say the hampered human relationship, followed by 46% saying the environmental pollution, 27% the economic imbalances and 23% each the misuse of drugs, adverse media impacts and increased warfare.

Such societal problems and harmful effects do emerge due to the inherent contradictions that exist between the competitive but cooperative nature of the progress of productive S&T worldwide on the one hand, and the monopoly appropriation of the gains on the other. This does not permit one to give more thrust to prevent harmful effects. So, one has to resolve this contradiction by harmonizing production relations and not resist S&T progress as such.

S&T: Its substantial contributions

Public perception on certain areas in which S&T could make substantial contributions is conditioned by their direct experiences. It is more logical but a simplistic view. About 60–80% of respondents assert that S&T could make substantial contributions to certain areas such as increased production and productivity of goods and services, exploration of space, earth and ocean and optimum utilization of resources, better, fast, safe and economical means of transportation and communication and better and cost-effective child birth control methods and devices. About 50–60% respondents think that S & T could make substantial contributions in providing better, effective and economic learning systems, improving the availability of quality food, control and prevention of environmental degradation and ecological imbalances, complement labour productivity, help better understanding of the laws that govern the development of nature, society and human thought, and reduce labour intensity and substitute labour force in production of goods and services. Still there are about 30–40% respondents who think that S&T could not make any substantial contributions in reducing natural disasters, in crime control and global wars by increasing international cooperation, in promotion of global trade and promotion of democratization of society by competitive and cooperative ventures. An equal proportion of the respondents express nil opinion.

On the other hand, crime is naturally controlled when the material and knowledge base of the society is democratized. That is, when people's entitlement to material and knowledge base is enhanced according to their capabilities and contributions, a more civilized and humane society is evolved. Such an entitlement is related to compulsory and universalization of education and employment. Reduction of threats from wars, in turn, is conditioned by the globalization of capabilities through competitive and cooperative spirits. This demands for promotion of global peace, understanding, harmony and respect for human values and rights. This, in turn, is conditioned by the knowledge and material base of the world people in general.

Such interdependences and interconnections among S&T and multiple human dimensions of world people are still in the meta-physical (abstract) state of development. It is due to the widespread disparities in capabilities among the world people. Hence, the need for correcting the negative and indifferent public perceptions and attitudes by proper education and mass awareness campaigns.

S&T areas: Public funding

Public perception on public funding in S&T areas is conditioned by the objective realities of the current societal issues. The public perception correlates to the thrust of research areas the country should focus while

allocating funds. About 70–80% of respondents assert that public funding in S&T should be towards areas like improving health care and population control systems, prevention of environmental pollution, ecological imbalances and strengthening of agriculture and food. About 50–70% stress on improving education, increasing economic productivity, developing better and cost-effective means of transportation and communication, clean and reliable sources of energy, remote sensing and weather forecasting, cost-effective and ethnically adaptive housing, and exploration of space, ocean and earth for optimal use of resources. However, a large section of respondents, 20–40%, express nil opinion. Neutrality of such a large section of population towards such topical issues could often be hurdles to the initiation of large-scale S&T projects which are otherwise needed for fulfilling the basic needs of the people. And India should increase its S&T expenditure, on accelerated basis, equal to or greater than 2% of GNP or 4 times the remittances for imported technologies, to reach the threshold level of technological assimilation and diffusion capability[4].

S&T fields: India in the lead

India was one of the ancient civilizations, whose contributions to world knowledge base, since Indus Valley Civilization, are well known to all of us. India was a centre of learning which gave birth to many philosophical connotations that changed the social dimensions of many countries in Asia and elsewhere. Rise of Buddhism is one such example. Such was our tradition till the dawn of the Middle Ages of Feudalism and British Colonialism. Though there was a setback to this tradition since industrial revolution in Europe and its colonial expansion, the Indian brains still possess creative talents, which are now latent and which can blossom any time given the right fertile conditions. Thus, relatively favourable conditions in the form of vast S&T, educational and industrial infrastructure created since independence has helped India to regain the lost vigour. This is reflected in the assertion of India's S&T capabilities in many key sectors. Thus, 52% of the respondents assert that there certainly are a few S&T fields in which India is in the lead. And 10% of the respondents assert that India is in the lead in most of the S&T fields while 28% assert none of the fields and 10% express nil opinion. When asked to name some such fields in which India is in the lead, the following in the order, have emerged.

1. Space research
2. Agricultural technology
3. Textile technology
4. Atomic energy
5. Software development and advanced computer programming
6. Mathematical and social sciences and philosophy
7. Medicine and health care research.

Interestingly, these are the same S&T fields which have come under international pressures in the form of Missile Technology Control Regime (MTCR), Nuclear Nonproliferation Treaty (NPT), Trade Retaliations, Super and Special 301, and Intellectual Property Rights (IPR), Trade-Related Intellectual Property Rights (TRIPS), Trade-Related Investment Measures (TRIMS) in the Dunkel Draft Text (DDT) of General Agreement on Tariffs and Trade (GATT).

Despite such pressures, India cannot reverse its historical course of advancement. This can be better done by subjecting India's international relations, including trade and technology transfer, to the five principles of peaceful coexistence.

S&T impacts: Life-style changes

Life-styles are culture-specific. Culture is a function of man's economic relationship with nature on the one hand and other fellow human beings in the society on the other in the process of production, distribution and consumption of goods and services. The economic relationship, in turn, is a function of S&T knowledge, that induces changes in values, attitudes and aptitudes towards environment that surround the people. S&T also provides material base in the form of better goods and services that ultimately brings desirable changes in the life-styles. Accordingly, the public is well aware of the impact of S&T in changing the life-styles. Thus, about 44% respondents feel that the rate of change of life-styles due to S&T progress is too fast while 38% think neither fast nor slow, 10% feel too slow and 8% express nil opinion. Telecommunications and electronics, transportation, automation, food and agriculture, information technology, domestic appliances, medicine, computers and engineering sciences are some S&T systems, which, according to the respondents, cause the rate of change of life-styles too fast. In fact, S&T progress enables democratization of the knowledge base worldwide, leading to cultural cross-fertilization of thought processes that ultimately weed out the obsolete customs and traditions. Once such hurdles are broken, S&T becomes a material force bringing rapid changes in values and life-styles. Information technology plays a vital role in bringing such changes. And such changes, however slow or fast, are irreversible. Therefore, suitable educational programmes should be developed to enlighten the public to acquire positive attitudes towards life-styles.

Part II: Development issues and action plans

We have posed many currently debated development issues before the public and sought their perceptions,

and action plans that could guide our country's S&T programmes towards emerging world trends. The consensus outcome, as presented below, is more precise and objective, hence worthy of consideration before undertaking S&T-oriented large-scale investment ventures.

Intellectual Property Rights

Intellectual Property Rights (IPR) in the form of patents, trademarks and copyrights are granted by the country's constitutional due process as a legal instrument to promote public good, while at the same time, protecting the legitimate interests of the first inventor. Any invention which is novel and possesses economic applications is considered for grant of patent rights. But products derived from chemical, biological, biochemical and microbial reactions shall not fall under the categories of inventions or innovations as their outcomes are inherent to the processes or methods. Likewise, life forms are the products of biological evolutions independent of man's efforts. Therefore, what is inventive or innovative here are the particular processes or methods developed by men and hence worthy of granting patent rights. On the other hand, a strong IPR regime, granting patents indiscriminately and kept in force beyond the economic utility of the inventions, becomes an instrument of obstructions in the hands of monopolies, which restrict the advancement of productive forces by competitive spirits[5]. Therefore, a desirable IPR regime is one which is liberal and sensible, and promotes productivity by competitiveness and respects self-reliance of nation states. Indian Patent Act 1970 is one such, primarily aimed at promoting technological innovations, transfer and commercial production in the country. And thus it was applauded as a model patent law by all the developing countries. Even this was the case with the patent laws of the developed countries in their early stages of development. But, in recent times, India has been subjected to international pressures, trade sanctions and retaliations such as Super and Special 301, TRIPS, TRIMS of DDT of GATT, to alter its Patent Laws fundamentally to fall in the line with the strong patent regimes. This international pressure is brought under the belief that a strong patent regime indeed will help promote global trade and commerce. The Parliamentary Standing Committee on Commerce, in its 3rd Report on Dunkel Draft proposals, presented in December 1993, urged the government not to succumb but to move with caution.

In this context, about 86% of the respondents did not accept the demand for strong IPR (Patent) regime, nor did they want India to sign DDT unless GATT agrees to revise its text to accommodate the legitimate economic interests of developing countries, including India. Only 14% of the respondents favoured the demand for strong IPR (Patent) regime and signing the DDT as it is presented today. A similar trend can be seen in the write-ups of as many as 20 lead authors that appeared in the April special issue of *JSIR*[6] that devoted itself to IPR and development perspectives. Here about 65% of the authors favour defending the Indian Patent Act 1970 and country-specific liberal patent regimes, while 35% of the authors favour changing the Indian Patent Act 1970 to align with the strong patent regimes of Paris convention and DDT. Still there is a need for enlightening the public about the science of patent laws, their inherent limitations to protect the monopoly market rights of the first inventors, as in the modern S&T world; parallel discoveries, inventions and innovations can take place autonomously by the surging competitive productive forces. Under these conditions, India has to conduct its foreign relations, including trade and technology transfer, subject to the five principles of peaceful coexistence, while at the same time, safeguarding the basic features of the Indian Patent Act 1970. Now India has been a signatory to DDT of GATT (WTO). And people should learn to get along with it while safeguarding national interests as well. And the Parliament is divided on the Bill to amend the Indian Patent Act 1970, the majority not favouring.

Nuclear Nonproliferation Treaty and Global Security Plan

Nuclear Nonproliferation Treaty and Global Security Plan are the schemes operated by the clubs of 'nuclear-haves', mostly the West, to control the spread of nuclear weapon systems among the 'nuclear-havenots'. This does not discriminate between nuclear programme for peace and energy and for warfare. And there is an international pressure on India to join this exclusive club of NPT. Or else it will have to face sanctions and retaliations on imports and exports in strategic areas of nuclear energy research. Despite this, the 25 years of NPT has led to manifold proliferation of the weapon[7]. By the end of this century, there will be 20 more nations in addition to 26 with full capacity to develop their own nuclear weapons[8].

In this context, 76% of the respondents did not want India to accept NPT membership while 24% favoured it. Further, they made certain startling but relevant assertions indicating their perceptions and attitudes. Thus, NPT is not a treaty reached among equals but an imposed one. GSP is irrelevant as there is no external threat to this world. This means that the threat to this world security is from within. NPT and GSP can be accepted if there is a universal agreement to dismantle the nuclear warheads of all the countries simultaneously. Even if India joins NPT and GSP, it should not be

under duress. Similar views were expressed by a group of eminent persons in the first quarter of 1993 commissioned by the Government of India (GOI) to look into this aspect. Even a high-level study group, commissioned by the Carnegie Endowment for International Peace, in its report entitled 'India and America after cold war' suggests US not to force India to sign NPT and MTCR under duress. And the majority of the respondents think that India should reduce its budget on the Department of Atomic Energy by making nuclear power generation a commercially viable and self-supporting venture. One can reduce the burden of defence by improving friendly relationship with neighbouring countries through improved diplomacy. This demand is nothing but a call for the renewal of India's one-time cherished foreign policy based on the five principles of peaceful coexistence and nonalignment, the *Panch Sheel*: Mutual respect for each other's territorial integrity and sovereignty; mutual nonaggression; mutual noninterference in each other's internal affairs; equality and mutual benefit; and peaceful coexistence.

Earth summit and biodiversity

Concern to save the earth, protect the environment, ecology and human habitat and preserve biodiversity (in fact, it should be rightly called as *biovariety*) has attained international dimensions in the recent times. The developed countries were the forefathers, who created those systems that caused environmental pollution and ecological imbalances. But that was the way the industrialization took place under the capitalist system. The same countries, after 200 years of capitalist industrial development, exploiting world resources at their will have now attained a higher level of human development index (HDI). As S&T progresses, people's concern for a clean environment increases. The developed countries could now come with third- or fourth-generation technologies, mostly automated, which are more environment-friendly. So, the sealing was fixed on the upper limit to environmental pollution to 1992 level, which the developed countries would like to impose on the developing ones. If the developing countries, with scarce resources, technologies and skills, want to reach the same level of development of developed ones via the routes of capitalist industrialization, they have to invariably depend on the inefficient, obsolete and environment-unfriendly technologies. Such obsolete technologies, now facing a threat of closure in the West, can be easily shifted to the developing countries. Therefore, one group in the West is eager to shift such obsolete units to willing partners in the developing countries. At the same time there is another group in the West which wants to put restrictions on the development as well under some pretext with which some other groups in the developing countries would like to align either due to necessity or due to ignorance.

Likewise, the developed countries would like to consider the biodiversity (biovariety) of the world (mostly from developing countries) as universal properties and hence the demand for free access. At the same time, any improvement they make on them through S&T inputs they would like to treat as private properties, protected by IPR regimes, which the developing countries will have to acquire by paying a high price.

These are the contradictions of development inherent in the collective nature of production of knowledge and materials and the private nature of appropriation of the surpluses thus generated. The UNO (United Nations Organization) wished to reconcile the conflicts among these contradictions and strike a balance under the existing diverse sociopolitical–economic systems. This drama was enacted on the stage of the Earth Summit in June 1992 at Rio. In this context, an overwhelming majority (84% of the respondents) did agree that the Earth Summit and Biodiversity Convention 1992 were good beginnings while 16% did not consider so. However, it is observed here that the Rio Treaty and Biodiversity Convention were heavily biased towards Multinational Corporations (MNCs), which according to the majority of respondents should be checked. According to the respondents, country-specific action plans to protect national interests should be incorporated into the Treaty. Mass awareness campaigns and peace movements should be taken up vigorously to restore the lost varieties of life forms in the high tides of green revolution. India should formulate suitable policies to protect biodiversity on cooperative basis and strengthen environmental protection policies and Acts while venturing into large-scale industrial developments.

S&T in the strategic areas: Missile Technology Control Regimes

Like NPT, Missile Technology Control Regime (MTCR) is the club of advanced countries who have stock-piled heaps of missile systems. Members of MTCR are presumed to have freedom of access to high-techs in the strategic areas. But they are barred from entering into business contracts with non-MTCR member countries, mostly the developing ones, to pass on the missile technologies. Thus, by restricting missile proliferation, MTCR member countries believe that they could safeguard the global security.

As Indian space research ventured to launch Polar Satellite Launch Vehicles (PSLV), Geostationary Satellite Launch Vehicles (GSLV) programmes with cryogenic rocket engine systems, obtained from Russia on economy terms, the US has imposed a ban, under MTCR, on ISRO, restricting its deals with US companies in strategic

and aerospace industries under the pretext that Indian space programme is aimed at developing missile technologies. However, if India joins MTCR and accepts to pay higher prices, the US itself is willing to pass on the cryogenic rocket engine systems to Indian space programmes. This, according to the Chief of Glavkosmos, is nothing but a trade war[9].

In the context of trade-related conflicting interest, as cited above, over 86% of the respondents did not favour India accepting MTCR terms. They assert so for the reasons that MTCR is an imposed one and that it does not help those who want to acquire high-tech systems. The ways suggested by the respondents are to strengthen peace movements that prevail over the arms race. Further, according to the respondents, India should use its missile technology development programmes to strengthen high-tech areas in the country and then extend its cooperation among the Third-World countries.

S&T-oriented large-scale investment projects and the priorities

Today, India is torn between antagonistic group interests. Different groups think that they are exclusive to one another in the development process. Ignoring the fact that the human development is an integral part of ecological balance, the multiple dimensions of social development are counterposed to one another, instead of enabling them to play complementary roles. Thus, we see, in India, that agriculture is counterposed to industries and large-scale industries are counterposed to small-scale industries, foreign trade is counterposed to domestic trade, domestic R&D and technologies are counterposed to import of foreign technologies, large-scale S&T-oriented investment projects in irrigation, power, space, defence, communication, transport, industry and infrastructure are counterposed to environment, ecology and human habitat, population itself is counterposed to development process and so on and so forth. Even modern S&T is counterposed to the future of human civilization. Thus, we see a stagnation (or rightly stafglation) in the development.

Such conflicts emerge out of the objective realities that India continues to live with the backward and obsolete productive forces sustained by backward and obsolete production relations which were the legacies of feudal and colonial rules. However, there is a surge of technologically driven productive forces advancing in the world simultaneously raising the living standards of the people by changing their production relations on more or less democratic norms. Indian society, as an integral part of the world, is also influenced by this trend, hence the urge for advancing the productive forces. But the groups desiring to advance the productive forces through noncontractual vested interests ignore the simultaneous need for restructuring a more balanced and harmonious production relation. This is the source of conflict that counterposes several dimensions of human development, resulting in stafglation. Public perception, in this study, is more objective and balanced one. They understand the development needs of the state and the country by advancing productive forces while at the same time they appreciate the demand for restructuring production relations so as to avoid the conflicts of counterposing nature. Thus, over 74% of the respondents desire to give top priority to S&T-oriented large-scale investment projects as they give opportunities to raise the living standards of the people. Yet they assert that such large-scale investment projects often induce negative impacts on environment, ecology and human habitats. The basic principles to be respected here are the laws of conservation of biomass and bioenergy that enhance the net worth of all the groups concerned in the development process. In the absence of such corrective measures, the public prefers a large number of small projects. They assign top priority to the projects aimed at eradication of poverty, unemployment and illiteracy. They urge for reduction of import dependence in defence sectors as such a dependence does not guarantee national security. Further, they urge that defence sectors should be made to serve civil needs as well. The respondents assert that satellite communication systems, hydroelectric power projects and nonconventional energy systems should be given higher priority. The public urges not to convert natural forests into biological parks. According to them, priority to nuclear power plants can be given if the safety is guaranteed and power generation is made to earn its own return on investment. Further, they suggest to accord priority to missile technology programmes only as a means to advance high-techs in the strategic fields and earn return on investment by trading in defence systems and hardwares.

Agriculture, industry, food, health care and family welfare programmes and the priorities

Agriculture and industry depend on each other for continued growth. Agriculture being the base, industry must play a complementary role as a superstructure. India, being an agricultural country, should not ignore this fact. The public perception to this that agriculture, purely on commercial grounds, is going to endanger the food security hence needs to be checked. Educated unemployed should be urged to take up vocations in agriculture and industry. All wastelands should be brought under agriculture and awareness should be created to cultivate every piece of land. Development of agriculture and industry should promote employment and economy. And S&T should play an active role in promoting agriculture, industry, employment and

economy. Mechanization of agriculture and automation of industry should enhance labour productivity and not replace the labour force.

Food security to the people is the national security. A threat to food security to the people is a threat to national security as well. Likewise, health is wealth. Unhealthy people means unhealthy nation, and hence insecure nation. Family welfare programmes should not only aim at regulating population growth in accordance with the means to feed but should also ensure food and health to one and all. There is a unity of consensus of all the respondents that top priority should be given to these programmes. The public asserts that more investment should be allocated to such areas. And S&T should focus more to meet the need for 225–250 million tonnes of food grains for 1044 million people in 2001 AD by raising the yield per hectare manifold from the present level of: rice 1.75 tonnes, wheat 2.27 tonnes, jawar 0.82 tonnes, maize 1.52 tonnes, ragi 1.1 tonnes, potato 16.22 tonnes and sugar 5.55 tonnes[10]. It is here that the productivity of foodgrains in the country, i.e. the output per work force, should be enhanced manifold by maximizing the product of two factor ratios, viz. output per hectare of land and hectare of land per work force. Maximization of the output per hectare demands for higher orders of productive technologies and maximization of hectare of land per work force demands for implementation of Land Reform Act as well as sharing of scarce resources by cooperative spirits. Simultaneously, population explosion should be controlled by persuation and education following the Kerala model. Decentralized administration and acceleration of green revolution are suggested. Stringent checks on adulteration of food and drinks are sought. Indigenous health care systems to be strengthened on the lines of modern health care systems. One should avoid mixing up of the politics of religion, castes and communal divides with the politics of development.

Education and employment

Education and employment are the two sides of the same coin. Education is useless if it does not lead one to gainful and productive employment. Most of the country's antagonistic contradictions could be resolved if compulsory and universal education and employment are ensured to all eligible and able-bodied people. Compulsory and universal education and employment to all meet the twin objectives of people's entitlement to the wealth of knowledge and the wealth of materials on more democratic and socialized basis. In this sense, Kerala has achieved a near cent percent success in universal education. And it has yet to go a long way in ensuring compulsory and universal employment to all. Whereas India, as a whole, is many years behind in achieving both these objectives. As per the 1991 census, the literacy rate was 52.2% with a wide gap between males (43.88%) and females (39.42%). During the 1981–91 decade, the growth rate for literacy was 8.55% and that for illiteracy was 7.31%, leaving 302 million illiterates even in 1991. And about 63% of the population in the age group of 14–60 years is considered as nonworking though the unemployed registered with the employment exchanges was as high as 36.3 million in 1991, which grows at the rate of 10% per annum[11].

The public perception, in this study, desires to give encouragement to community level action plans and vocationalization of plus-two stage for attaining education and employment to all. The purpose of education should be to build a vibrant, productive self-confident, self-reliant, scientific-tempered, value-oriented, more humane and cultured society. S&T should help to create such a society.

Housing and human habitat

Housing and human habitat, in harmony with environment and ecology, is the index of developed and cultured civilizations. According to 1991 census, the demand for housing in Kerala was 11 lakh units[12]. Of these, 55% were substandard huts declared unfit for human dwelling, 27% were old ones awaiting demolition and reconstruction; only 1% were considered as standard units fit for human dwelling and 17% were considered as numerical shortage. A similar trend exists on all-India basis as well. Thus, housing the millions is a massive public task. If such a massive task is taken in a coordinated manner, it could provide immense opportunities for investment, employment, innovative and productive techniques, methods and materials. The criterion on the choice of technique is the one that delivers maximum number of housing units per given unit of resources.

The public perception here is more positive. Accordingly, S&T should focus on cost-effective, modular-type, functional and economy houses by adopting Baker's type of techniques which encourage utilization of local resources and skills that conserve materials, reduce costs and improve aesthetic values in harmony with environment and ecology. Housing design and construction should be fully engineered based on architectural principles rather than on false prestiges. Minimization of noncontractual vested interests will also reduce the costs while improving cost-effectiveness.

National defence

Agriculture, industry, S&T and defence are considered as four vital elements of any nation. Without strengthening these four aspects concurrently, one cannot think

of a modern state. They are complementary to each other and not counterposing. If improperly handled and imbalances are created, they do become counterposing to one another. Though the public perception (about 85% of the respondents) is in favour of strengthening national defence, their attitude is conditioned by the imbalances that occurred due to the stifled nature of development as seen in agriculture, industry, S&T and defence. Thus, there is a demand for cut-down in defence expenditure, unless otherwise the nation wants to do business in the military hardware. Expenditure in defence can be brought down by strengthening peace movements, imparting military training to the civilians to develop para-defence systems and improving country's diplomacy and foreign relations.

1. *The Public Understanding of Science*, The Royal Society, London, 1985.
2. Zhongliang, Z., *Sci. Public Policy*, 1991, pp. 311–317.
3. Qin, L. Y., *Nature*, 1993, **365**, 200.
4. Govindarajulu, V., *J. Sci. Ind. Res.*, 1983, **42**(9), 477–483.
5. Daniele, A., *Sci. Public Policy*, 1992, **19**, 357–368.
6. *J. Sci. Ind. Res.*, 1993, **52**(4), 209–330.
7. *The Hindu*, 15 April 1995, p. 14.
8. *The Plain Truth*, 1987, **52**(1), 12–163.
9. Vladimir, R., *The Hindu*, 7 April 1995, p. 14.
10. *The Hindu Survey of Indian Agriculture*, 1992, pp. 29–81.
11. *Manorama Year Book*, 1993, pp. 531–533.
12. Tampi, K. M., *The Hindu*, 25 October 1993, p. 5.

ACKNOWLEDGEMENTS. I am grateful to Kerala State Committee for Science, Technology and Environment for the research grant. I thank all those enthusiastic people who responded to this study and my colleagues for their kind cooperation.

Impact of extreme events on coastal zones and small islands in the context of climate change

S. M. Kulshrestha

The paper presents a review of the current state of knowledge on the likely impacts of extreme weather and climate events on coastal zones and small islands in the context of climate change. Possible response options, including prevention, preparedness and adaptation, are indicated.

AT first glance, there appears to be little in common between the gradual onset and complex impacts of climate change and the unheralded onset and direct impacts of extreme events. While the magnitudes of the anticipated global warming and consequent sea level rise may be somewhat less than those indicated in the First Assessment Report[1] of the Intergovernmental Panel on Climate Change (IPCC), it is the extreme events and/or climate surprises that need to be taken seriously. It appears that more emphasis has to be placed on changes in frequency and/or intensity of various extreme or threshold events, particularly the extremes of temperature and precipitation, and the hazardous coastal events such as cyclones and storm surges. Such events would dominate the effects in coastal zones. The small islands would, however, appear to be highly vulnerable to accelerated sea level rise as well.

Definitions

Coastal zone: Coastal zone includes both the area of land subject to marine influence and the area of sea subject to land influence. Coastal zones are very valuable parts of the economies – both the affluent and subsistence economies. Coastal habitats provide important areas for fish and wildlife, including many endangered species. Coastal zones filter and process agricultural and industrial wastes and provide a buffer for the inland areas against cyclone and wave damages.

Small islands: Islands less than 200 km^2 in area are defined as small islands and those with area less than 20 km^2 may be called very small islands. This categorization is not entirely arbitrary and is significant in the sense that while the larger islands generally have geological features and problems similar to those found in continental regions, the small and very small islands have their own characteristics and rather unique problems. The major conglomerates of small islands are in the Pacific, the Atlantic (Caribbean), the Indian Ocean, and the territories of Indonesia and the Philippines.

Physical changes

Tegart *et al.*[2,3] have documented the potential impacts

S. M. Kulshreshta resides at B-10/7233, Vasant Kunj, New Delhi 110 070, India.

of climate change based on the studies conducted for preparing the 1990 Assessment[1] and the 1992 Supplement[2] Reports for IPCC. Among the likely physical changes in coastal zones and small islands as a result of possible climate change, the potential for, and magnitude of, accelerated sea level rise (ASLR) appears to be the greatest, especially for small islands. The projected ASLR (of the order of half a metre by the year 2100) would result in (i) substantial land inundation with especially significant impacts on coral atolls, (ii) considerable loss in agricultural activity in low-lying areas of coastal zones and small islands, (iii) coastal inundation and erosion resulting in degradation and loss of natural protective features such as dunes and mangroves apart from property damages, (iv) increased backwater effects reducing drainage from coastal streams, and (v) severe damage to surface and ground water sources due to increased intrusions of salinity. Nevertheless, as mentioned above, the major changes would be due to extreme events and climate surprises, which would also result in sudden rise of sea level apart from the other physical changes.

Accelerated sea level rise (ASLR)

Estimates of the combined effects of oceanic thermal expansion and land ice changes that could be induced by global warming involve considerable uncertainty[5]. These uncertainties are: (i) the rates of greenhouse gases (GHGs) emissions and the related changes in atmospheric concentrations and radiative forcings, (ii) feedbacks in the climate system as they affect the climate sensitivity to a given change in radiative forcing, (iii) transfer of heat from the surface to deeper waters in the oceans, (iv) regional and seasonal changes in climate, (v) model uncertainties in the computed response of ice sheets and glaciers to climate changes, and (vi) changes in ocean circulations for which adequate models are yet to be validated. Among these, thermal expansion appears to be the most significant long-term change[6]. It needs to be recognized at this stage that the ongoing research may, in the near future, result in new projections for the magnitude and the consequent physical changes due to ASLR resulting from climate change.

Enhanced extreme events – cyclones, storm surges, and heavy rains

The IPCC 1992 Supplement[4] cautioned that confidence in regional climate change patterns based directly on global climate model (GCM) output was low and that there was no consistent evidence of change in the variability of storminess. The position remains the same at present. In order to tackle this uncertainty, the models have to refine the horizontal resolution apart from other improvements.

Evans[7] made a comprehensive study of the envisaged impacts of enhanced greenhouse warming on tropical cyclones with special reference to the Australian region and concluded that in the overall, according to the present GCM simulations, preferred cyclogenesis locations remain unchanged but the potential maximum intensity could increase with the caveat that the environmental effects are ultimately the limiting factors on actual cyclone intensity and that motion will remain largely unchanged. Later work, for example, by Raper[8] indicates that there is no convincing evidence to support a relationship between sea surface temperatures (SSTs) and cyclone intensities and that an increase in SSTs may not necessarily lead to increase in the intensity of cyclones. As regards the frequency of cyclones, the link with SST appears to be through the prevailing atmospheric conditions to which both the SSTs and the tropical cyclone frequencies are closely related. Although SSTs could increase in the future due to climate change, the model predictions indicate only small changes in circulation, usually within the range of the past natural variability. There is some indication for certain areas (like the USA mainland) that the number of hurricanes striking the coast in recent years is a bit less than the average number based on the statistics of the past 100 years or so.

While evaluations of the possible impacts on coastal and estuarine surge levels due to climate change involve complex site-specific analyses, the impacts of possible ASLR on storm surge levels can be assessed on the basis that the storm surge height above the sea level plane for a given wind shear stress will be lower with higher sea level positions, where an off shore shelf is characterized by a horizontal sea bed. In the more common situations of sloping shallow offshore shelves and varying basin configurations, no direct general relationships can be deduced in terms of sea level positions and the relative differences of surge heights for given wind forces[9].

As regards heavy rain events, various lines of evidence suggest an increase under enhanced greenhouse conditions. This will probably be accompanied by an increase in deep convection and a decrease in more widespread light rain and number of rainy days in middle latitudes. These changes, coupled with increased evaporation, would result in significant reductions in the return period of floods as well as droughts[10].

Snow and ice accumulation and ice outflow

It is generally accepted that as the cause of sea level rise, the thermal expansion of oceans will dominate over the other two factors, viz. the melting of nonpolar glaciers in places like the Himalayas and New Zealand and changes in the volume of ice and snow in Antarctica

and Greenland. Practically all the GCMs suggest that an increase in precipitation over Antarctica will occur with global warming[11]. This should lead to some increase in the rate of accumulation of snow as has been observed over the recent decades in a study of ice cores[12]. The contribution of this to changes in sea level is small. Nevertheless, the complex question of snow and ice accumulation and ice outflow in Antarctica is important to the long-term question of sea level rise because the West Antarctic ice sheet, being marine-based, is potentially unstable and vulnerable to climate change. If it were to disintegrate, it would contribute about 6 m to sea level rise but such an event would occur only over many centuries[13]. Even then, this possibility is considered sufficiently significant and major research projects have been taken up[14]. Ice sheet instabilities cannot explain all aspects of the long-term climate dynamics and ocean instability needs to be taken into account but it is the ice sheet instability that seems to trigger ocean instability.

Lightning from violent thunderstorms

Lightning strikes from violent thunderstorms are very hazardous to electric power installations, aircraft, forest and the people. Even when widespread fire does not occur, lightning can cause serious damage to forests, thereby enhancing the spread of insects and diseases apart from the economic loss. According to a recent study[15], initial model results show that for a doubling of carbon dioxide scenario, global lightning activity would increase by approximately 25%. Among the climatic variables that have an impact on dangers of wildfires, relative humidity is the most important. This underlines the need for site-specific studies of the relative impacts due to increased lightning and the likely changes in relative humidity regimes locally.

Ecological impacts

The various global scenarios of sea level rise indicate an accelerating rise with no prospect of stabilization. If the seas continue to rise, coastal erosion will also accelerate and become more widespread as any compensating sedimentation declines. It will encounter, reshape, and in due course submerge the raised beaches. Because of the complexities of coastal environments and their along-shore variations, a great deal of sector-specific research is still needed to assess the extent of the geomorphological and ecological changes that will occur as the sea level rises. In assessing the nature and extent of possible changes, it is useful to classify coasts into several categories such as steep and cliffed coasts, beach fringed coasts, deltaic coasts, swampy coasts, estuaries and coastal lagoons, intertidal and near-shore areas, coral reefs and reef islands, and man-made coasts. An excellent discussion of the impacts of the rising sea level in coastal regions is given by Bird[16].

Socioeconomic impacts

In coastal zones and small islands, as indeed elsewhere as well, the socioeconomic impacts of climate change are both on the local environment and on the human societies. Those on the local environment can be analysed in terms of (a) ice and snow, (b) oceans and coasts, (c) hydrological cycle, (d) natural hazards, and (e) ecosystems and vegetation. Similarly, the possible impacts of climate change on the local human societies can be assessed in terms of (a) water resources, (b) food and agriculture, (c) coastal and island dwellers, (d) health and (e) economic activity. In actuality, the impacts occur in more complex ways and impacts in one sector affect/create further impacts in other sectors. Coastal zones and small islands are likely to suffer erosion, coastal flooding, inundation, agricultural losses, and damages to water resources and other essential infrastructure. Ironically, certain impacts could lead to a clash of interest. For example, as a result of rise in ground level temperatures due to global warming, beach seasons in coastal or island resorts would be extended bringing in more revenues – probably much more than the costs of controlling erosion and inundation that may be caused by sea level rise on account of climate change. Higher SSTs may enable new swimming areas to be developed where sea water is at present too cold to swim. On the other hand, in certain cases, adverse impacts would get accentuated such as flooding from inland waters and storm surges and rising sea level from the ocean side.

Response options – prevention, preparedness, insurance

Prevention

It is debatable if the potential impacts of climate change can be 'prevented'. Appropriate adaptation or abatement efforts are possible and should be encouraged although cost–benefit considerations will apply especially in countries with weak economies. Reduction in the emission of GHGs (including methane from garbage dumps), energy conservation, adoption of cleaner energy and manufacturing technologies, and control of atmospheric pollution (especially that from automobile exhausts) are a few of such strategies that are widely recommended.

As regards wildfires or bushfires, one technique for prevention is to light small fires (called hazard reduction burns) that consume the fuel so that during subsequent

bushfire, there is less fuel and hence less likelihood of a large uncontrollable fire[17].

Preparedness

It is possible to reduce the losses significantly if the existing knowledge and the known adjustment systems are employed[18]. These authors have listed a wide variety of available tools as follows:

Long-term
- Controlling investment, planning and development
- Management of land, resources and infrastructure
- Education and training
- Public information
- Hazard insurance
- Weather modification
- Structural engineering works
- Hazard-resistant constructions

Immediate pre-impact
- Monitoring
- Predicting
- Warning
- Shelters
- Evacuation and relocation

Impact and immediate post-impact
- Emergency management
- Hazard-fighting
- Search, rescue and salvage
- Disaster relief
- Loss compensation and rehabilitation

Insurance

The international insurance industry is already faced with a significant increase in insurance outflow due to natural disasters/extreme events during the last few decades. There are increasing demands on the insurance industry owing to the increasing population and value-added property. This is much more apparent for coastal zones, especially in affluent countries and in tourist areas, although insurance still does not command much presence in small islands except where high-value property and infrastructure are involved. The present problems faced by the insurance industry would increase if climate change results in an increase in frequency or intensity of extreme events. Meanwhile the available scenarios and projections of climate change and its impacts would tend to increase insurance premia rates, especially in coastal zones and small islands. Reinsurance would become a major and regular practice in such areas because of the magnitude of economic risk being very high and beyond the capacity of one insurer.

Adaptation options

General options

If strategies to limit the emission of GHGs are successful, the impacts on coastal zones and small islands would also be reduced. However, even if the limitation strategies succeed, there will be impacts due to climate changes already set in motion. Moreover, it is likely that initially limitation would be partial at best. Consequently there are some urgent requirements form coastal adaptation, which can be considered in the from of the following activities:

(a) Improved understanding of the climate change science and likely implications for sea level rise and other impacts such as cyclones.
(b) Monitoring of sea level and coastal change.
(c) Vulnerability studies to identify areas most likely to be prone to extreme events and sea level rise.
(d) Site-specific impact assessments.
(e) Forward planning.
(f) Public awareness.
(g) Development of a full range of coastal adaptations, including nonstructural or nonengineering options.

Specific options

In specific terms, the adaptation options for coastal zones and small islands fall into three categories, viz.

(a) Retreat: Abandon vulnerable areas and resettle.
(b) Accommodate: Adjust to the impacts, including ASLR.
(c) Protect: the vulnerable areas, especially those with large populations, high-value property and precious natural resources, and strategic installations.

An excellent discussion of the available adaptation options is given in the Report of the Coastal Zone Management Sub-Group of IPCC Working Group III[19] which also lists a very large number of references on the subject.

Social and cultural implications

People inhabiting coastal zones and small islands have distinct life-styles with characteristic social and cultural ethos. The adaptation option should not unduly strain their social and cultural values.

Sensitivities and thresholds

Definitive prediction of future climate seems impossible because a number of uncertainties are involved. Some can be reduced but others, especially those related to

future human behaviour, climate surprises and catastrophic geological or geophysical events, cannot be anticipated adequately. As future climate cannot be predicted satisfactorily (at present in any case), assessing the possible impacts of climate change requires the use of sensitivity studies over a range of possible values for relevant climatic variables. This approach enables response surfaces to be plotted showing the value of the property or the outcome considered as a function of a range of possible values of the climatic variables. Response surfaces are particularly useful because the effect of any chosen climate change scenario can be read off the graph without the need for new calculations. Examples of response surfaces are shown by Whetton *et al.*[11].

Another approach could be through calculation of quantitative indices that combine the state of climate and the economic activity into a useful number which may be regarded as an 'econoclimatic' indicator.

Conclusions

Coastal zones and small islands are very vulnerable to climate change, especially changes in extreme weather and climate events. The confidence in regional climate change patterns deduced from global climate models (GCMs) is low at present and considerable progress will have to be made before the output products of GCMs can be applied at regional or local scales. Hence, there is need for site-specific assessments of the likely impacts of climate change (especially those of extreme events)in coastal zones and small islands so that locally applicable response strategies and policy options could be developed.

1. Houghton, J. T., Jenkins, G. J. and Ephraums, J. J. (eds), Scientific Assessment of Climate Change, Report prepared for IPCC by working Group I, Cambridge University Press, 1990.
2. Tegart, W. J. McG., Sheldon, G. W. and Griffiths, D. C., Climate Change, The IPCC Impacts Assessment, Report prepared for IPCC Working Group II, Australian Government Publishing Service, 1990.
3. Tegart, W. J. McG. and Sheldon, G. W. (eds), Climate Change 1992, The supplementary Report to the IPCC Impacts Assessment, Australian Government Publishing Service, 1993.
4. Houghton, J. T., Callandar, B. A. and Varney, S. K. (eds), Climate Change, Supplementary Report to the IPCC Scientific Assessment, Cambridge University Press, 1992.
5. Warrick, R. A., in *Climate and Sea Level Change* (eds Warrick, R. A., Barrow, E. M. and Wigley, T. M.), Cambridge University Press, 1993.
6. Wigley, T. M. L. and Raper, S. C. B., *Nature*, **357**, 1992.
7. Evans, J. L., Envisaged Impacts of Enhanced Greenhouse Warming on Tropical Cyclones in the Australian Region, Technical Report No. 20, Division of Atmospheric Research, CSIRO, Australia, 1990.
8. Raper, S. C. B., in *Climate and Sea Level Change* (eds Warrick, R. A., Barrow, E. M. and Wigley, T. M.), Cambridge University Press, 1993.
9. Flather, R. A. and Khandker, R., in *Climate and Sea Level Change* (eds Warrick, R. A., Barrow, E. M. and Wigley, T. M.), Cambridge University Press, 1993.
10. Pittock, A. B., Fowler, A. M. and Whetton, P. H., Proceedings of the International Hydrology and Water Resources Symposium, Perth, 1991.
11. Whetton, P. H., Hennessy, K. J., Pittock, A. B., Fowler, A. M. and Mitchell, C. D. (eds), Regional Impact of the Enhanced Greenhouse Effect on Victoria, Annual Report, 1991–92, Climate Impact Group, Division of Atmospheric Research, CSIRO, Australia, 1992.
12. Morgan, V. I., Goodwin, I. D., Etheridge, D. M. and Wookey, C. W., *Nature*, 1991, **354**, 58–60.
13. Budd, W. F., *Climate Change*, 1991, **18**, 271–299.
14. Bindschadler, R. A. (ed.), *West Antarctic Ice Sheet Initiative, Science and Implementation Plan*, vol. 1, NASA, 1991.
15. Price, C., Proceedings of the Symposium on the World at Risk: Natural Hazards and Climate Change, Center for Global Change Science, Massachusetts Institute of Technology, 1992.
16. Bird, E. C. F., *Submerging Coasts – The Effects of a Rising Sea Level on Coastal Environments*, John Wiley, 1993.
17. Beer, T., Allan, R. and Ryan, B., in *Modelling Climate Change in Environmental Systems* (eds Jakeman, A. J., Beck, M. B. and McAleer, M. J.), John Wiley, 1993.
18. Mitchell, J. K. and Ericksen, N. J., in *Effects of Climate Change on Weather-related Disasters in Confronting Climate Change: Risks, Implications and Responses* (ed. Mintzer, I. M.), Cambridge University Press, 1992.
19. Misdorp, R., Dronkers, J. and Spradley, J. R. (eds), Strategies for Adaptation to Sea Level Rise, Report of the Coastal Zone Management Sub-Group to the IPCC Response Strategies Working Group, Ministry of Transport and Public Works, The Netherlands, 1990.

New pressure-induced phase changes in CdWO4 from Raman spectroscopic and optical microscopic studies

A. Jayaraman, S. Y. Wang and S. K. Sharma

University of Hawai'i at Manoa, Hawai'i Institute of Geophysics and Planetology, Honolulu HI 96822, USA

Pressure-induced phase transitions in ABO_4 tungstates and molybdates with scheelite and wolframite structure are of interest to materials science and mineral physics. High-pressure phases with interesting physical properties can be expected. Through a high-pressure Raman and optical microscopic study we have found two novel pressure-induced phase transitions in wolframite-type $CdWO_4$. The first transition near 10 GPa is subtle, but easily observable under the polarizing microscope, as the crystal breaks into domains of colour, presumably due to a phase change accompanied by twinning, or ferroelasticity. At the second transition near 20 GPa, the Raman data clearly indicate a change in W–O coordination from basically tetrahedral to truly octahedral, a new finding of significance to high-pressure crystal chemistry.

PRESSURE-induced phase transformation in scheelite (tetragonal symmetry) ABO_4 compounds have been of current interest from the viewpoint of their high-pressure crystal chemistry[1]. Recent high-pressure Raman spectroscopic studies[2,3] have revealed that pressure-induced phase transformations are a rule in these systems. Another common structure adopted by ABO_4 tungstates at ambient pressure is the so-called wolframite type[4] and to our knowledge these have not been studied under high pressure.

Two questions of great interest in the high-pressure behaviour of ABO_4 molybdates, tungstates and rehenates are (i) whether the scheelite-type compound would transform under pressure to the wolframite-type structure and (ii) how stable is the wolframite structure itself with respect to high pressure. The best candidates to study in this connection are $CdMoO_4$ and $CdWO_4$, because the former crystallizes in the scheelite structure and sits on the border line between the scheelite–wolframite stability field[4], and $CdWO_4$ has the wolframite structure and hence is a very good candidate for clarifying the second question. Therefore, we carried out high-pressure Raman spectroscopic studies on $CdWO_4$ and $CdMoO_4$ up to 40 GPa and optical microscopic studies on $CdWO_4$ up to 20 GPa. In this article we will discuss mainly our Raman results on $CdWO_4$, and argue for a change in the coordination of tungsten to a truly octahedral coordination with respect to oxygens at the pressure-induced phase transition observed near 20 GPa. We will briefly discuss our results on $CdMoO_4$, which undergoes the anticipated scheelite to wolframite transformation as well as the phase transition involving the above-mentioned coordination change. Finally, we will present optical microscopic evidence for a subtle phase change near 10 GPa in $CdWO_4$.

High-pressure Raman experiments on $CdWO_4$ were carried out in a gasketed Mao-Bell-type diamond cell, with 4 : 1 methanol–ethanol mixture as the pressure-transmitting medium. Pressure was calibrated by the well-known ruby fluorescence technique. Raman spectra were generally recorded using a Spex Raman instrument equipped with photomultiplier. In some experiments a Spex Triplemate instrument provided with a CCD diode array detector was used to record the spectra. Figure 1 shows the Raman spectra of $CdWO_4$ obtained at three different pressures. The bottom spectrum belongs to the wolframite phase, the middle one has Raman peaks belonging to a new high-pressure phase plus the wolframite phase, and the top spectrum represents purely the high-pressure phase. Figure 2 *a* and *b* show a plot of the observed Raman peak positions as a function of pressure. Near 20 GPa the spectrum clearly changes due to a pressure-induced phase transition. Raman peaks of the wolframite phase extend beyond 20 GPa and disappear completely only above 30 GPa. Likewise, on release of pressure the high-pressure phase persists down to about 16 GPa. Below this pressure the sample is purely that of the wolframite phase. The appearance of the two-phase region shows very clearly that the transition is strongly first-order and that it is hysteretic. However, the transition is completely reversible at room temperature and the high-pressure phase is not quenchable. Now the question is what is the nature of the high-pressure phase? Only a high-pressure X-ray structure investigation can give a decisive answer, but the Raman results contain some important information regarding the structural building block, namely the coordination of the W–O polyhedra.

Detailed structural investigations on the scheelite-type ABO_4 compounds have shown that the B cation is tetrahedraly coordinated to oxygens, while the A cations

are surrounded by eight oxygens[4,5]. Raman and infrared spectra[6–8] are consistent with the presence of isolated BO_4 tetrahedra in the structure[7]. In ABO_4 of the wolframite type, such as $CdWO_4$, the B cation is regarded as sixfold-coordinated (approximately, octahedrally coordinated) with respect to oxygens, while the A cation is similarly situated with six oxygen neighbours[4]. The structure may be crudely described as hexagonally close-packed oxygens with certain octahedral sites filled with A and B cations in an ordered way[4].

The wolframite-type $CdWO_4$ has monoclinic symmetry and belongs to the space group I2/a (ref. 9). The usual view of the structural arrangement is that it consists of zigzag chains of edge-sharing WO_6 octahedra. But the description of the building blocks as octahedrally coordinated WO_6 is a highly distorted picture, for the two of the W–O distances are much larger than the four other distances to regard them as regular WO_6 units. For instance, in the isomorphous $NiWO_4$ these distances are found to be 1.79 Å for four W–O distances and 2.19 Å

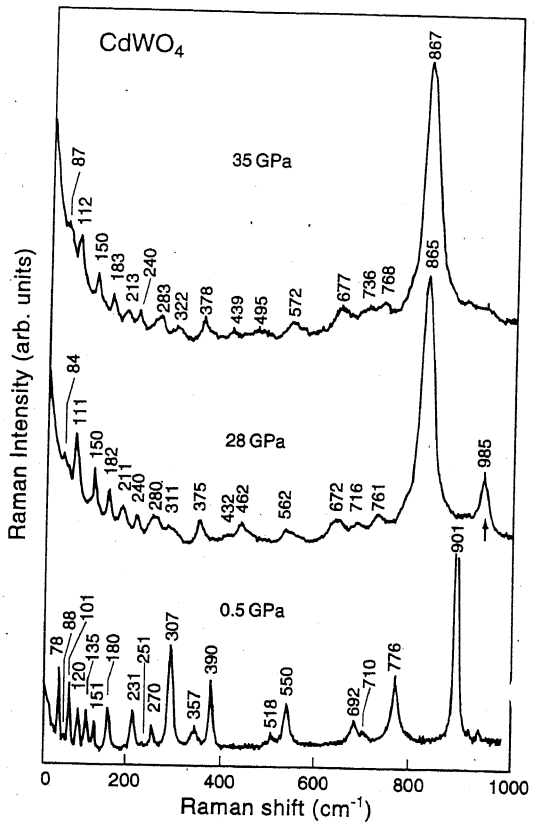


Figure 1. Raman spectra of $CdWO_4$ recorded at three different pressures with 514.5 nm line of the argon ion laser. The spectrum at 0.5 GPa is of the wolframite phase, which is the normal phase. The 28 GPa spectrum has changed and the strongest line representing the ν_1 mode frequency (865 cm^{-1}) has dropped by 120 cm^{-1} from that of the wolframite phase (985 cm^{-1}), seen to the right. The sample has not fully converted to the new phase. The wolframite peak is marked by an arrow. The spectrum at 35 GPa represents purely the new phase (see text for explanation).

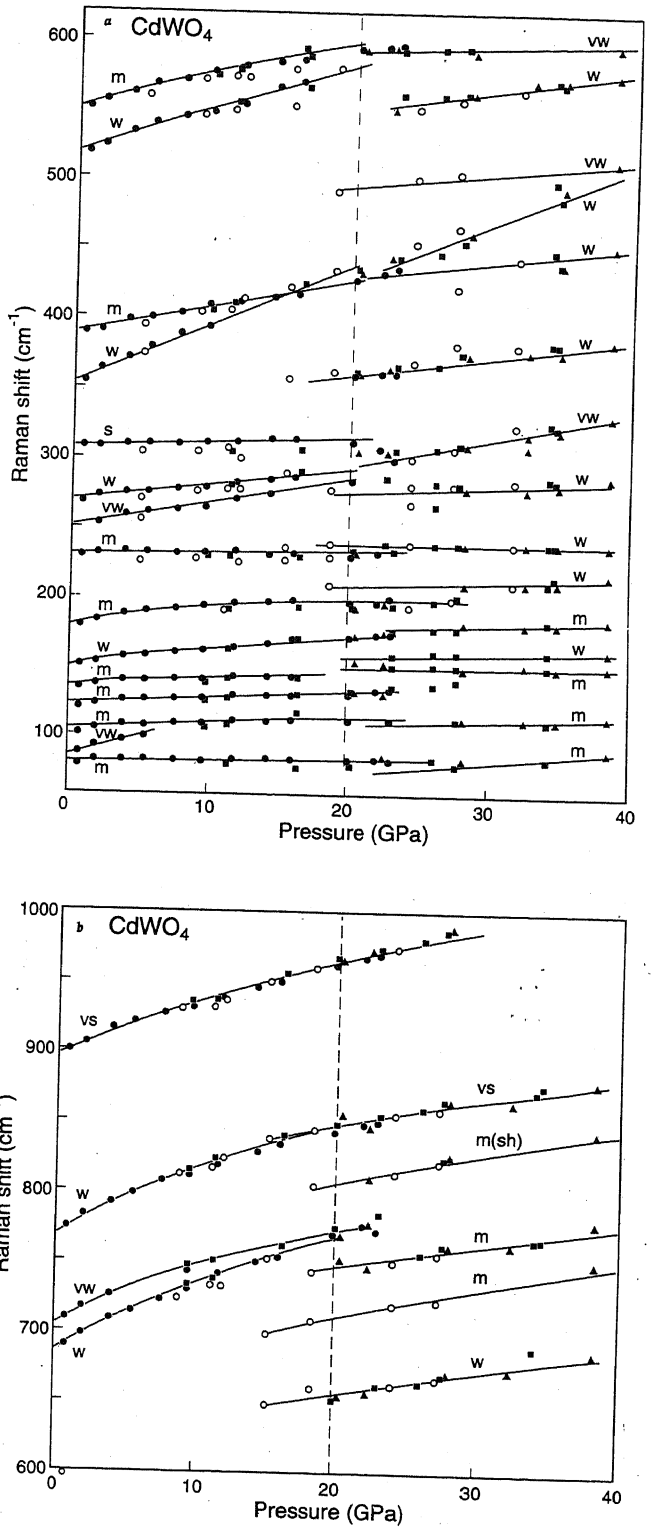


Figure 2 *a* and *b*. Pressure dependence of the observed Raman peaks in the wolframite phase as well as in the new high-pressure phase. A phase transition is marked by sudden shifts in the Raman peak frequency as well as appearance of new peaks and disappearance of old peaks. The dotted line marks the transformation pressure. The symbols s, m, w and vw appearing in the figure indicate the intensities of the Raman peaks, strong, medium, weak and very weak, respectively. Sh is shoulder. Solid lines have been drawn through the data points.

for the other two[10]. Further, it has been pointed out[7] that the Raman frequencies of WO_4 in wolframite-type $CdWO_4$ are consistent if the W–O coordination is regarded as tetrahedral rather than octahedral.

The remarkable feature in the Raman spectrum of the high-pressure phase is that the ν_1 mode frequency, which represents the totally symmetric W–O stretching vibration, drops by about 120 cm^{-1} from that of the wolframite phase at the corresponding pressure (see Figure 1) and dominates the spectrum. This is precisely what would be expected if the W–O coordination becomes truly octahedral. When the coordination increases, the bond lengths increase, and the result is a drop in mode frequency. For instance, the ν_1 mode frequency of the octahedral tungstate group (WO_6) in Ba_2CaWO_6 (refs 11, 12) crystallizing in the ordered perovskite structure is about 100 cm^{-1} lower compared to that of WO_4 in tetrahedral coordination[7]. Comparing our results with the above we conclude that in the high-pressure phase of $CdWO_4$ the coordination for W becomes truly octahedral. This would require that the WO_6 octahedrons form chains in the structure sharing four corners with neighbouring octahedra, or form zigzag chains sharing two edges, to preserve the W–O ratio. The Raman spectrum has other weak peaks in the high-frequency region which must be identified with the other internal modes appropriate for WO_6 as well as with the external modes. Obviously, the crystal symmetry will decide the splittings of both the internal and external modes. Input from a high-pressure X-ray diffraction study is needed to identify the symmetry of the high-pressure phase, which will then enable one to assign the other observed peaks. But the identification of the highest wave number peak with the ν_1 mode of the WO_6 seems reasonable, and this is a valuable piece of information that we get from Raman spectroscopy, regarding the structure of the building block. Further, the change in colour of the sample after the 20 GPa transition would also be consistent with increase in coordination, for now the crystal field acting on the 5d states of W would be much larger, resulting in an enhanced 10 DQ splitting. This would lower the conduction band and reduce the optical gap, causing the observed colour change.

Our high-pressure Raman studies on $CdMoO_4$ (ref. 13) show two pressure-induced phase transitions, one near 12 GPa and another near 25 GPa. From a detailed comparison of the Raman features of the first high-pressure phase with that of wolframite-type $CdWO_4$ it is clear that the first pressure-induced phase transition in $CdMoO_4$ is from scheelite to wolframite type. In the second transition near 25 GPa the behaviour is very similar to that observed in $CdWO_4$. The ν_1 mode drops by about 120 cm^{-1} and the spectrum is similar to that of the high-pressure phase of $CdWO_4$. This indicates that the wolframite phase transforms to a structure in which the Mo–O coordination becomes octahedral. Thus, the results of $CdMoO_4$ and $CdWO_4$ are consistent.

We have made optical observations under a microscope on thick (~ 50 μm) as well as thin (~ 20 μm) samples of $CdWO_4$ in polarized and unpolarized white light. Remarkably, a pressure-induced phase transition that is not evident in the Raman results could easily be observed near 10 GPa. The sample is normally colourless in unpolarized white light and has a uniform colour in crossed polarization. The onset of the phase transition is marked by the sudden appearance of light grey bands inside the crystal when viewed in unpolarized light. In crossed polarization the effect is spectacular, for the sample transforms into parallel domains of different colours with sharp boundaries. These are illustrated in Figure 3 *a–c*.

In crossed polarization thin samples show the colour domain vividly. With increasing pressure numerous grey-coloured parallel bands develop in the sample when viewed in unpolarized light, which in crossed polarization has a colour pattern corresponding to the bands in Figure 3 *c*. At the 20 GPa transition the pattern breaks up and is replaced by dark spaghetti-like regions, the sample acquiring a yellow colour. The yellow progressively changes to a deep orange colour with increasing pressure. Figure 3 *d* illustrates the appearance of a thicker sample near 40 GPa as seen by transmission under white light illumination.

When pressure is released, the sample does not recover its original appearance, once it breaks up into domains. The grey bands are still seen and the originally observed uniform colour over the sample area in crossed polarization is not restored. After the 20 GPa transition a reversal to a single crystal is not expected and this is borne out from our observations.

The optical observations are strong evidence for a pressure-induced phase transition near 10 GPa. In fact, when the Raman data are examined critically, a few minor changes are seen, viz. small changes of slope in the pressure vs. Raman shift plot and a small (4 wave numbers) abrupt downward shift of the strong Raman peak near 304 cm^{-1}. Now the interesting question is: What is this subtle transition in the wolframite phase? A strong possibility is that the wolframite lattice I2/a takes on the C2/c symmetry of $HgWO_4$. This would involve only a reorientation of the doubled unit cell of wolframite by a rotation around the *b*-axis, as shown by Jeitschko and Sleight[5] for the reverse case, namely the relationship between wolframite structure and the structure of $HgMoO_4$. In this transition the new unit cell volume would be the same, but the angle β is expected to be much larger than for wolframite $CdWO_4$, ~ 115° vs. 91.5°. Jeitschko and Sleight[5] describe the differences between the two structures as follows. In both the structures the oxygens form close-packed layers

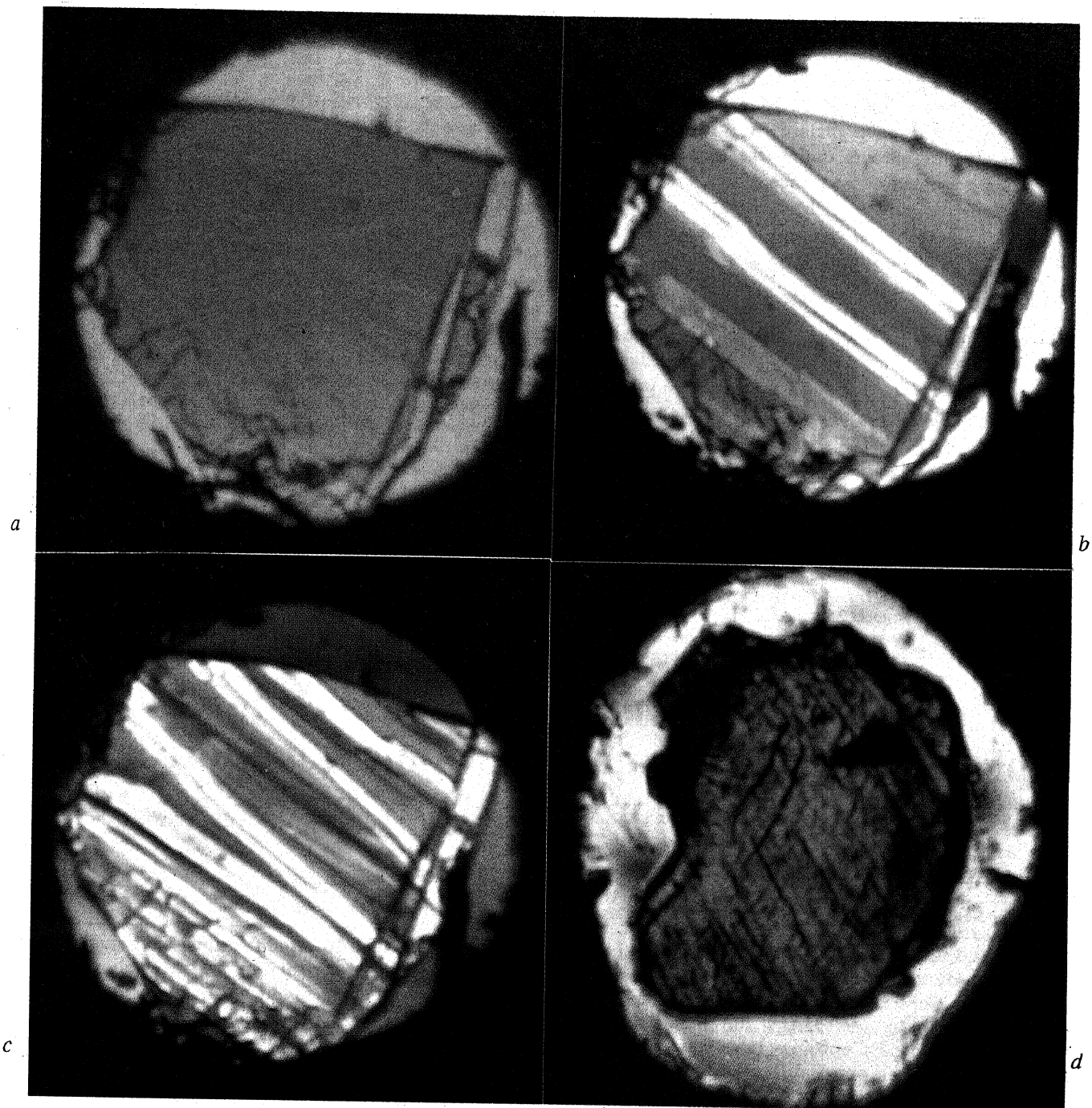


Figure 3. Photomicrographs (magnification × 400) of two $CdWO_4$ samples in the form a thin plate: *a*, *b*, *c*, taken with white light in crossed polarization at about 0.5 GPa, 9.7 GPa and 11 GPa, respectively. The uniform pink colour indicates that the sample is single crystal. In *b* the colour domains are due to a subtle phase transition (see text for explanation). In *c* the pressure has been increased and the result is that the domains multiply. The domains could be due to twinning or ferroelastic nature of the new phase. *d*, indicates the appearance of a thicker sample in transmission, with white light (unpolarized) as the source, at a pressure of about 38 GPa. The deep orange colour is believed to be connected with the second pressure-induced phase transition which involves a change in the W–O coordination, from basically tetrahedral to truly octahedral coordination. The latter transition occurs near 20 GPa and the colour of the sample progressively turns from yellow to deep orange. The sample is practically colourless and transparent before this phase transition.

parallel to the *b*–*c* plane. However, the stacking of these layers is almost perfectly hexagonal close-packed in the $CdWO_4$ wolframite structure, while it would be close to cubic close-packing in $HgMoO_4$, and $HgWO_4$. The wolframite-type $CdWO_4$ adopting the $HgWO_4$ structure would be consistent with the trends in high-pressure crystal chemistry, namely that the structure of the compound down the column in the Periodic Table is a likely structure for the high-pressure phase of the analogous compound just above it in the same column. Single-crystal high-pressure X-ray diffraction on $CdWO_4$ would be very rewarding in this connection and it is quite

feasible in the present case, since the transition pressure is 10 GPa, within the reach of hydrostatic pressure.

The subtle nature of the transition as reflected in the Raman results and the observed optical property changes would fit in the above picture. It is very likely that multiple twinning occurs during the transition and the domains are a direct consequence of it. Twinning would also explain the fact that the sample does not regain its original single-crystal character on release of pressure after the 10 GPa transition. The other possibility is that they are ferroelastic domains. The optical observations under a polarizing microscope demonstrate the power of this technique in revealing subtle phase transitions.

Our picture then for the high-pressure transformation sequence in $A^{2+}B^{6+}O_4$ compounds is scheelite to wolframite, or wolframite-like phase and then to a phase with truly octahedrally coordinated BO_6 units. There are other monoclinic structures closely related to wolframite for ABO_4 compound[5] and they are a possibility as intermediate high-pressure phases, before the transformation to the phase involving a change to octahedral coordination.

1. Nicol, M. and Durana, J. F., *J. Chem. Phys.*, 1971, **54**, 1436–1440; Ganguly, N. and Nicol, M., *Phys. Status Solidi B*, 1977, **79**, 617–622.
2. Jayaraman, A., Batlogg, B. and van Uitert, L. G., *Phys. Rev.*, 1983, **B28**, 4774–4777; *ibid*, 1983, **B31**, 5423–5427; Jayaraman, A., Kouroklis, G. A., van Uitert, L. G., Grodkiewicz, W. H. and Maines, R. G., *Sr. Physica*, 1988, **A156**, 325–340; Jayaraman, A., Wang, S. Y., Shieh, S. R., Sharma, S. K. and Ming, L. C., *J. Raman Spectrosc.*, June 1995.
3. Christofilos, D., Kourouklis, G. A. and Ves, S. (to be published).
4. Sleight, A. W., *Acta Crystallogr.*, 1972, **B28**, 2899–2902.
5. Jeitschko, W. and Sleight, A. W., *Acta Crystallogr.*, 1973, **B29**, 869–875.
6. Tarte, P. and Liegeois Duychaerts, M., *Spectrochim. Acta*, 1972, **A28**, 2029–2036; Liegeois-Duykaerts, M. and Tarte, P., *ibid.*, 1972, **A28**, 2037–2051.
7. Blasse, G., *J. Inorg. Nucl. Chem.*, 1975, **37**, 97–99.
8. Clark, G. M. and Doyle, W. P., *Spectrochim. Acta*, 1966, **22**, 1441–1446.
9. Chichagov, A. P., Ilyukhin, V. V. and Belov, N. V., *Dokl. Akad. Nauk SSSR*, 1966, **166**, 87–89.
10. Keeling, R. O., *Acta Crystallogr.*, 1957, **10**, 209–213.
11. Corsmit, A. F., Hoefdraad, H. E. and Blasse, G., *J. Inorg. Nucl. Chem.*, 1972, **34**, 3401–3407.
12. Blasse, G. and Corsmit, A. F., *J. Solid State Chem.*, 1973, **6**, 513–518.
13. Jayaraman, A., Wang, S. Y. and Sharma, S. K. (to be published).

ACKNOWLEDGEMENT. Work supported by NSF, Material Science Division, Grant No. DMR-9402443.

Received 24 April 1995; accepted 9 May 1995

Analysis of a conformation-specific epitope of the alpha subunit of human chorionic gonadotropin: Study using monoclonal antibody probes

N. Venkatesh, G. Nagaraja and G. S. Murthy*

Centre for Reproductive Biology and Molecular Endocrinology, Indian Institute of Science, Bangalore 560 012, India

Monoclonal antibodies (MAbs) have been used extensively for identification of sequence-specific epitopes using either the ELISA or/and IRMA methods. However, attempts to use MAbs for identification of conformation-specific epitopes have been very few as they are considered very labile. We have investigated the stability of conformation-specific epitopes of human chorionic gonadotropin (hCG) using a quantitative solid-phase radioimmunoassay (SPRIA) technique. Several epitopes are stable to mild modification (chemical and proteolytic) conditions, and epitopes show differential stability for these modifications. Based on these observations, a monoclonal antibody (MAb 16) for an α-subunit-specific epitope of hCG has been used to monitor changes at the epitopic site (identified as epitope 16) on modification of hCG, using SPRIA with immobilized MAb 16. Modifications of amino groups, hydroxyl group of tyrosine as well as carboxyl group of Asp/Glu all bring about sufficient changes in the epitope integrity. Peptide bond hydrolysis at lysine residues damages the epitope, but not at arginine residues. Hydrolysis at tyrosine does not affect the epitope, though modification of the side-chain of tyrosine inactivates the epitope. Destruction of the epitope occurs on reduction of the disulphide bonds. Partial retention of the epitope activity is seen on modification of carboxyl or the ε-amino groups of lysine. Based on these results four to six amino acids have been identified to be at the epitopic site, and the data suggest that two peptide segments are brought together by the disulphide bond Cys10–Cys60 to form the epitope.

EPITOPES form a very important and specific immunochemical feature of a protein or a macromolecular antigen, and this feature of the antigen is used extensively in the development of specific radioimmunoassays, immunodiagnostics, vaccines etc.[1–4]. Epitopes are classified into sequence-specific (contiguous) and conformation-

*For correspondence.

specific (discontiguous, assembled, topographic) epitopes[5–10]. In fusion experiments MAbs are produced against both the sequence-specific and conformation-specific epitopes. Identification of these epitopes would reveal the immunochemical property of an antigen as well as provide useful clues to the three-dimensional structure of the antigen molecule. Current methods of analyses of sequence-specific epitopes using the synthetic peptides are popular and very widely used[11–17]. However, contradictory views exist as peptide fragments are poor substitutes for native antigen[18–25]. This may be because of the inability of the synthetic peptides to adopt the conformation similar to that of the native protein. Majority of antigenic determinants, which seem to consist of strongly conformation-dependent epitopes as revealed by X-ray crystallographic data and peptide binding studies[26–31], could not be uncovered by peptide approach. With the present state-of-the-art of peptide synthesis it is difficult to alter the conformation of peptides synthetically. Yet a generalized approach is not seriously attempted for the identification of conformation-specific epitopes, even though a need for it is felt acutely[32,33]. We have recently developed a quantitative solid-phase radioimmunoassay (SPRIA) technique using MAbs which extends the utility of low-affinity MAbs as well as MAb culture fluids to use them as quantitative probes for assessment of epitope integrity. Using this analytical method we demonstrate that in hCG several conformation-specific epitopes are differentially stable to treatments like enzyme digestion, chemical modification, etc., and utilize the MAb SPRIA to identify the amino acid residues contributing to the integrity of the α-specific epitope 16.

Materials and methods

Human chorionic gonadotropin (hCG) was immunoaffinity-purified from early pregnancy urine. SDS–polyacrylamide gel electrophoretic pattern was identical to purified CR 125 preparation, NIH. Iodogen method[34] was used to iodinate hCG to a specific activity of 50 μCi/μg. All assays were performed using MAb culture supernatants generously provided by Rajan Dighe, CRBME, IISc. The production and characteraction of the MAbs have been described previously[35]. $Na^{125}I$ (carrier-free) was purchased from Amersham, UK. All other reagents were of analytical grade.

Preparation of MAb-coated tubes

High-density polyethylene tubes were coated with an immunochemical bridge using a procedure similar to that described earlier for rabbit system[36]. These tubes were washed with RIA buffer, coated with appropriate dilution of MAb, and left at room temperature overnight. The following day the MAb solution was poured out, washed twice with RIA buffer and stored at 4°C until use.

SPRIA

MAb-coated tubes were incubated with 100 μl of ^{125}IhCG (100,000 cpm) and serially diluted standard hCG or modified hCG in a total volume of 600 μl of RIA buffer for 20 h at room temperature. On the subsequent day the contents of the tubes were discarded, washed twice with RIA buffer and counted in an LKB multigamma counter. ^{125}IhLH was used as a tracer in SPRIA for MAb 3 because of its better binding.

Chemical modifications of hCG

(a) Acetylation was done using acetic anhydride according to the established procedure[37]. Briefly, 200 μg of hCG in 60 μl of half-saturated sodium acetate was reacted with a total of 2 μl of acetic anhydride taken in saturated sodium acetate and vortexed, and the reaction allowed to occur for 10 min at 0°C. Acetic anhydride was added 4 times at intervals of 10 min for completion of the reaction. At the end the reaction mixture was diluted to 1 ml with water and stored for further reactions/assays at −20°C.

(b) Trinitrobenzene sulphonic acid (TNBS) reaction was carried out in 1% sodium bicarbonate solution at hCG concentration of 5 mg/ml and TNBS concentration of 0.15%, at 37–40°C for 2 h (ref. 38). The reaction was terminated by addition of 1 mg glycine in 50 μl water. The solution was made up to 1 ml in water and stored at −20°C.

(c) Biotinylation was carried out by the standard procedure using *N*-hydroxysuccinimidobiotin[39]. 150 μl of 4 mg/ml hCG in 1% sodium bicarbonate solution was mixed with 150 μl of a solution of *N*-hydroxysuccinimidobiotin (4 mg/ml in dimethyl formamide, DMF) overnight at room temperature. hCG with DMF was incubated as control. The reaction was terminated by adding 5 mg lysine and the solution was made up to 1 ml with water and stored at −20°C.

(d) Glutaraldehyde–diaminoethane reaction was carried out by reacting 200 μg of hCG (200 μl water) with 2 μl of glutaraldehyde taken in 100 μl water, and incubating at room temperature for 2 h. The reaction was terminated by adding 100 μl of 5% ethanolamine/diethylamine/glycine at pH 7 and diluted to 1 ml with water and stored at −20°C until use.

(e) Tyrosine modification was done using tetranitromethane (TNM) by the standard procedure[40]. 10 mg/ml hCG in 1% sodium bicarbonate (20 μl) was reacted with

5 µl of 1% TNM in absolute ethanol. The reaction was allowed to occur for 2 h at room temperature and terminated by diluting the reaction mixture to 1 ml with water containing 1 mg tyrosine in phosphate buffer, pH 7.0.

(f) Cleavage of methionine bonds was done by the standard procedure[41] using cyanogen bromide (CNBr). Suitable hCG control was also obtained by incubating hCG in 70% formic acid. After the reaction the solution was lyophilized and diluted with water and stored at −20°C.

(g) Modification of carboxyl groups was done using the procedure of Rice and Etzler[42]. Briefly, 50 µg hCG (20 µl) in phosphate buffer, pH 5.0, was mixed with 20 µl of 15% diaminoethane and reacted with 20 µl of freshly prepared 0.2 M 1-ethyl-3-(3-dimethyl-aminopropyl) carbodiimide (EDAC) in water. The reaction was terminated at the end of 3 h (room temperature) by adding 60 µl of 5% glycine, the volume was made up to 1 ml with water and stored at −20°C.

Enzymatic modification. Trypsin, chymotrypsin, carboxypeptidase A (CPase) and leucineaminopeptidase (LAPase) digestions were done by the standard procedure. All digestions were done at pH 8.0 in 0.2 M sodium bicarbonate solution at 37°C. The concentration of hCG used was 2 mg/ml while that of trypsin and chymotrypsin was 1 mg/ml each. CPase and LAPase were used at 10 µl/digestion. The digestion was terminated by diluting the mixture with RIA buffer containing 1% bovine serum, and stored frozen at −20°C until assay.

Chymotryptic digestion followed by CPase digestion. Chymotryptic digestion was done in duplicate as already described. The digestion was terminated by inactivating the enzyme with PMSF (1 mg). The tubes were lyophilized. One of the tubes was digested with 10 µl of CPAse A overnight at 37°C.

Estimation of the amino group. Amino groups in proteins were estimated[43] using glutamic acid as standard.

SDS–Polyacrylamide gel electrophoresis. SDS–PAGE was carried out using PhastGel gradient 8–25 and PhastGel SDS buffer strips from Pharmacia as per manufacturer's instructions. Samples were dissolved in buffer containing 10 mM Tris · HCl, 1 mM EDTA, 2.5% SDS with or without 5% β-mercaptoethanol, left in boiling water bath for 5 min and cooled, and a load of 500 ng/µl was applied. Silver staining was done using the standard procedure.

CNBr digestion and chymotryptic cleavage. CNBr digestion was done by the standard procedure and the hydrolysate was dried over NaOH in a vacuum desiccator. This was dissolved in phosphate buffer and digested with chymotrypsin as already described.

Circular dichroism (CD). CD spectra of hCG and its TNBS derivative were recorded between 190 and 250 nm at room temperature with 1 mm path length quartz cell, using a Jasco J-500A spectropolarimeter equipped with 501N data processor. Spectra were averaged over four accumulations. Protein concentrations were based on dry weight (0.5 mg/ml in 10 mM $NaHCO_3$) and the results reported as molar ellipticity values in units of deg cm^2/dmol.

Results

The extent of hydrolysis as assessed by quantitating the increase in amino groups in hCG (Table 1) after treatment with trypsin, chymotrypsin and subtilisin, and its effect on the integrity of several epitopes as measured by MAb SPRIA is shown in Table 2. In the case of subtilisin digest the number of peptide bonds hydrolysed is about 16 and results in 90% loss of binding activity as would be expected. The number of amino groups hydrolysed as measured by the TNBS method is as high as 11 (trypsin) and 4 (chymotrypsin) for digestion at 24°C and 12 (trypsin) and 8 (chymotrypsin) at 37°C. Thus, about 66% of the possible tryptic sites and 70% of the chymotryptic sites are cleaved at 37°C. In spite of this high percentage of peptide bond breakage, the structures of several epitopes are not affected (epitopes 3, 6, 16 and 21 for chymotryptic digestion and epitopes 3 and 21 for tryptic digestion, and epitope 3 for TNBS reaction) or only partially affected in all other cases (retention of activity to > 20%), with significant loss of activity occurring at epitopes 16, 21 and D11 for TNBS modification, demonstrating the rigidity of the conformation of these specific epitopes for certain modifications and their susceptibility to other modifications. It should be noted that the stability pattern is not the same for all MAb and each MAb shows its own characteristic pattern.

The SDS gel electrophoretic profile of the native hCG (Figure 1) shows two bands corresponding to

Table 1. Estimation of amino groups in hCG and proteolysed hCG

Enzyme used	Max. susceptible bonds	Amino groups measured	Peptide bonds hydrolysed
None	—	12	0
Trypsin	18	23*	11
		24**	12
Chymotrypsin	11	15*	4
		20**	8
Subtilisin	—	28**	16

*Digestion done at 24°C.
**Digestion done at 37°C.

Table 2. Binding activity of proteolysed/modified hCG to different MAbs as assessed by SPRIA

Enzyme used	No. of bonds hydrolysed	Activity (%) to				
		MAb3	MAb6	MAb16	MAb21	MAbD11
None	None	100	100	100	100	100
Trypsin	12	600	5	66	100	20
Chymotrypsin	8	600	90	90	25	20
hCG–TNBS	—	100	73	11	4	4
hCG–subtilisin	16	—	—	10	11	8

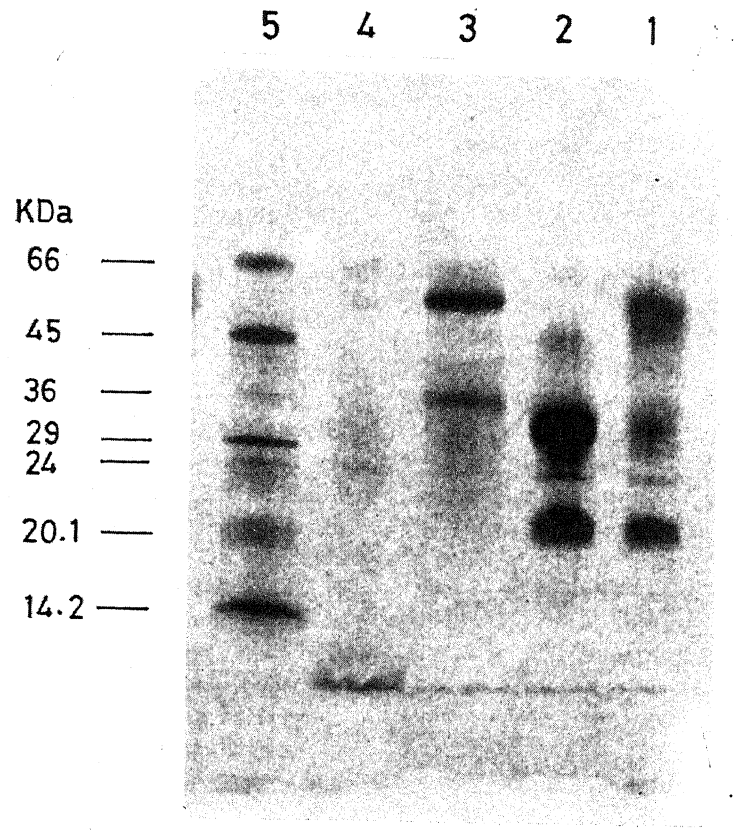


Figure 1. SDS–PAGE of hCG and its tryptic digest: lanes 1 & 2, native hCG under nonreducing and reducing conditions; lanes 3 & 4, tryptic-digested hCG under nonreducing and reducing conditions; lane 5, molecular weight markers (Sigma). Load was 500 ng in all cases and the bands were visualized by silver staining.

α and β subunits (lane 1) under reducing conditions. On digestion with trypsin these bands disappear completely, indicating that the large molecular subunits on proteolysis have been converted to small molecular peptides. In contrast, the digested hCG run under nonreducing conditions still shows large molecular subunit bands, indicating that the digested subunits are held together.

The effect of tryptic digestion (Figure 2, curve a) and acetylation (Figure 2, curve b) of hCG on its binding ability (immunoactivity) to MAb 16 shows a loss of 20% and 93% respectively. The extent of reduction of the binding activity (always refers to the ability of epitopic region to bind to MAb 16) of hCG on different chemical modifications shown in Table 3 indicates that all modifications of amino and carboxyl groups result in partial loss of activity from derivatives obtained at saturating concentration of the modifying reagent (Figure 3, where data are presented on the retention of activity with increasing concentration of TNBS used in the reaction). Complete loss of activity results only on modification of tyrosine by TNM. Neither formic acid nor peptide bond cleavage of methionine causes significant loss.

In contrast to these chemical modifications, enzymatic modifications resulted in no loss/marginal loss of activity

under normal conditions of digestion with trypsin, chymotrypsin, CPase and LAPase. However, digestion with trypsin at high concentrations (3 mg/ml) resulted in 80% loss of activity. Amino group modification of hCG with either acetylation or TNBS protected the epitope against trypsin digestion. While chymotrypsin or CPase individually did not affect the activity of the epitope, CPase digestion of the chymotryptic digest resulted in a loss of activity of about 30%. In another coupled digestion chymotryptic digestion of CNBr-treated hCG resulted in retention of activity of as much as 50%.

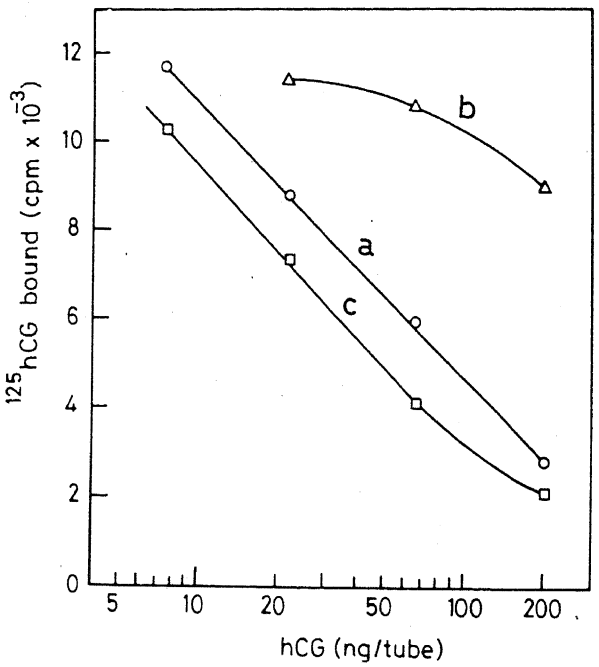


Figure 2. Displacement of ^{125}I-hCG from immobilized MAb 16 by control hCG (curve c), tryptic-digested hCG (curve a) and acetylated hCG (curve b). Tryptic digestion was done at a trypsin concentration of 1 mg/ml and hCG concentration of 2 mg/ml.

The CD spectra of hCG (Figure 4) were consistent with earlier reports[44,45] giving a negative CD band at 207 nm. The results of CD spectra of TNBS derivative indicate that no major conformational alterations have occurred due to modification with TNBS. Thus, the observed loss in epitope activity can be ascribed to the modification rather than any gross conformational change. This conclusion is further supported by the observation that all chemically modified derivatives exhibited retention of immunoactivity with one or other α-specific MAbs.

Identical results were obtained using modified derivatives of free α subunit (Table 4). As in the case of hCG, TNBS and TNM modifications resulted in loss of activity, while enzymatic modifications with trypsin and chymotrypsin resulted in partial/complete retention of activity. This observation clearly demonstrates that the presence of the β subunit is not blocking this epitope.

Discussion

Most of the studies on epitope mapping have been directed against sequence-specific epitopes, using peptide synthesis as a powerful tool in their recognition[11–17]. However, use of synthetic peptides has limitation in the studies on conformation-specific epitopes[18–25] as the conformation cannot be designed to that of native molecule by peptide synthesis. In addition, synthesizing disulphide-constrained peptides is rather difficult. Hence, attempts to use synthetic peptides have been limited, too laborious and expensive to be adopted routinely. Alternative ap-

Table 3. Effect of chemical modifications and enzymatic digestion of hCG on the binding activity to MAb16 as assessed by SPRIA

	Chemical modification			Enzyme digestion			
Reagent	Group modified	Percentage activity retained		Enzyme used	Amino acid hydrolysed	Percentage activity retained	
Ac_2O	Amino	7	5	Trypsin	K, R	10*	70**
TNBS	Amino	4	8	Chymotrypsin	F, Y	100	100
Biotinyl	Amino	3	—	CPase	C-term	100	—
GLD–DAE	Amino	10	8	LAPase	N-term	100	—
TNM	Tyr	< 1	—	AchCG + Tryp.	R	100#	100
HCOOH	—	100	—	CNBrhCG + Chymotrypsin	F, Y	50#	
CNBr	Meth.	100	—				
CDI/DAE	Glu/Asp	21	28	hCG–Chym + CPase A	Y65	70	

*Trypsin conc. of 3 mg/ml, 2.5 h at 37°C.
**Trypsin conc. of 1 mg/ml, 2.5 h at 37°C.
#The activity expressed is against acetyl hCG and CNBr–hCG taken as 100%, respectively.
All amino group modifications were complete, checked by TNBS.
All enzyme modifications resulted in increase of free amino group as measured by TNBS.
Ac, Acetyl; Gld, Glutaraldehyde; DAE, Diaminoethane; CDI, carbodilmide.

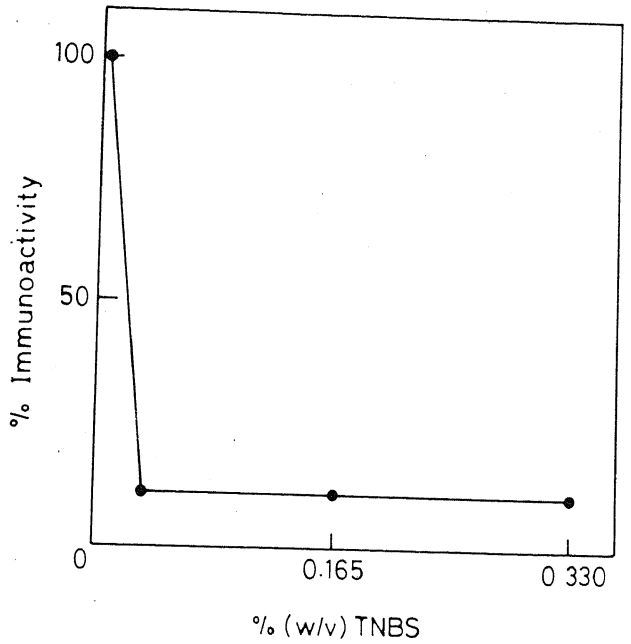


Figure 3. Inactivation of epitope 16 in hCG by varying concentrations of TNBS.

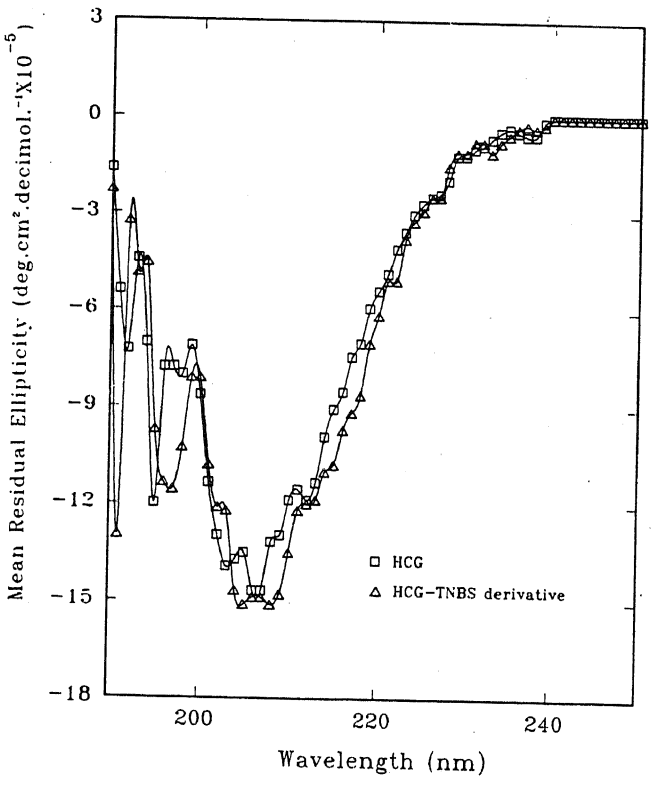


Figure 4. CD spectra of native hCG (-□-) and TNBS-modified derivative (-Δ-).

Table 4. Effect of modification of the α subunit of hCG on epitope 16

Modifying reagent	Group modified	Percentage activity retained
TNBS	ε amino groups	3%
TNM	Side-chain of Y	< 0.1
Trypsin	K and R peptide bonds	50
Chymotrypsin	Y and F peptide bonds	100

proaches for their study have not been attempted seriously as conformation-specific epitopes have been shown to be labile and inactivated easily[46]. The use of conventional approaches used in epitope mapping, namely ELISA, IRMA, etc., suffer from the disadvantage that they are mostly semiquantitative and limit the technique to high-affinity MAbs[47,48]. With the development of a MAb-SPRIA we were able to investigate quantitatively the effect of different types of modifications which do not destroy the native disulphide bonds on different epitopic regions, like proteolysis, chemical modification, etc. The results obtained indicate that several conformation-specific epitopes on hCG are stable to a variety of treatments and show the feasibility of development of a simple approach for identifying the amino acid residues at the epitopic region.

Proteolytic digestions of hCG have been carried out at very high concentrations of enzymes to ensure maximal proteolysis. Nonspecific proteolysis using subtilisin (Table 1) causes the hydrolysis of sufficient number of peptide bonds and results in more than 90% loss of structure in all epitopes investigated, as would be expected. Considering that the best of peptides to sequence-specific epitopes exhibit only about 1% cross-reactivity compared to their native structure in the protein, retention of activity up to a level of 20–100% after 4 h hydrolysis at 37°C with specific proteases like trypsin and chymotrypsin indicates the stability of several epitopes to proteolysis at Lys/Arg and Try/Phe residues. As seen from Table 2, four of the five epitopes used here are stable to either tryptic or chymotryptic proteolysis. It is certain that hydrolysis under these conditions has occurred by SDS gel electrophoretic pattern of the digest, where reduction of the disulphide bonds results in complete disappearance of the large molecular weight fragments, and amino group estimations of hydrolysed hCG (Table 1) using TNBS. In both the digests 65–80% of the susceptible bonds are hydrolysed. Yet, the loss of activity as seen by SPRIA is insignificant in most cases, and where marginal loss of activity is seen (like MAb D11 and 6 for chymotryptic digest), the retention of the activity is far superior compared to the best achieved with an identified peptide in sequence-specific epitopes.

The unusual stability of the epitope can possibly be attributed to the disulphide bonds. Each one of the epitopes loses more than 99% activity on reduction and

carboxymethylation. X-ray crystallographic structure of hCG published recently, also indicates that the disulphides are very closely clustered, conferring rigidity to the peptide chains in proximity[49].

The number of peptide bonds cleaved by trypsin under the conditions of hydrolysis is at least 12 of the possible 18, and yet the digest remains associated (Figure 1). This indicates the rigidity of heterodimer association, probably due to the extensive disulphide network. Based on these data, enzymatic hydrolysis of hCG results in only a superficial damage to the three-dimensional structure of the molecule as a whole. It would be interesting to know whether this observation is true for other proteins which do not have such an extensive disulphide bond reinforcement.

These results demonstrate the feasibility of proteolysis in conjunction with assessment of MAb activity as a powerful tool in identifying the amino acid residues contributing for the epitope conformation. TNBS modification of the amino groups of lysine residues results in a loss of about 90% activity (Table 2) in epitopes 6, 16 and 21, while the integrity of epitopes 3 and 6 seems to be unaltered. The gross structural integrity of the chemically modified hCG is evident from the CD spectra (Figure 4) and this attributes the losses of activity seen for epitopes 16, 21 and D11 to the presence/proximity of the ε amino groups of lysine residues in these epitopes. In contrast, the epitope regions 3 and 6 are not significantly altered by TNBS reaction, ruling out the proximity of the ε-amino group of lysines near these epitopes. Retention of epitope activity by TNBS modification of hCG in epitopes 3 and 6 clearly demonstrates that these epitopes do not have a side-chain residue of lysine nearby. This observation, namely different modifications affecting epitope integrity differentially, provides us with a tool which has been used in the analysis of epitope 16, described in the following.

Chemical modification of the amino groups of lysine residues done by several methods (Table 3) demonstrates that there is a significant loss of activity. While acetylation and glutaraldehyde–DAE modification reduces the activity by 90%, other reagents namely biotinylation and TNBS, reduce the activity by >95%. The CD data of the TNBS derivative (Figure 4) clearly show that gross structural changes are not occurring on modification. Thus, the loss of the activity of the epitope on modification may be ascribed to the presence of the ε amino group of a lysine residue at or near the epitopic region. Similarly, the side-chain of tyrosine is very closely associated with this epitope, as TNM modification results in almost total loss of activity. Here again the loss of activity cannot be ascribed to gross structural changes as the CD spectra of the TNM derivative rule out such a change. The presence of a carboxyl group near the epitope is also obvious as chemical modification of carboxyl groups reduces the activity to nearly 25%. These data clearly put a lysine, a tyrosine and an acidic amino acid at the epitope region. Reduction and carboxymethylation of the disulphide bonds of hCG completely abolish the activity (>99.9%) indicating the importance of the disulphide bridge in the integrity of the epitope.

Enzymatic hydrolysis of the hCG/derivatives (Table 3) indicates an important role for the peptide bond of a lysine or an arginine residue as hydrolysis by trypsin at low concentrations eliminates the activity significantly (30%), and to a great extent at high concentrations of trypsin (90%). This observation indicates that this bond is not easily susceptible to enzyme digestion, as has been shown earlier[50]. Tryptic digestion of acetylated hCG does not result in further loss of activity, and hence rules out the importance of arginine residue in the integrity of the epitope (Table 3). Similarly, a role for the peptide bonds of Y or F is ruled out as chymotryptic hydrolysis does not cause loss of activity. Retention of full activity by both LAPase digestion and CPase A digestion puts the terminal portions of the molecule away from the epitopic site. Hydrolysis of the peptide bond adjacent to methionine residues is also not important, since the CNBr-treated hCG retains complete activity.

The results of chemical modification of the α subunit also lead to the same conclusions. Proximity of a K and a Y side-chain is clearly indicated (Table 4). Similarly, the involvement of a lysine carboxyl peptide bond is obvious. Thus, all these results establish that the presence of the β subunit has not altered the epitope region.

PRIMARY SEQUENCE OF hCG ALPHA SUBUNIT

```
                    9 10
1  A P D V Q D C P E C T L Q E N P F F S Q

21 P G A P I L Q C M G C C F S R A Y P T P

                                      60
41 L R S K K T M L V Q K N V T S E S T C C

   61  63  65
61 V A K S Y N R V T V M G G F K V E N H T

                 88  90  92
81 A C H C S T C Y Y H K S.
```

Modification data indicate that the epitopic region is comprised of residues of lysine, tyrosine and an acidic amino acid nearer to a disulphide bond. Primary structure of the α subunit has one sequence having all the three amino acids in close proximity (Y88–S92). CPase hydrolysis does not result in loss of the epitope activity and hence rules out the contribution of these residues to the epitopic structure. However, there are other sequences C60–Y65 and C10–E9, which bring the three

amino acids to proximity through the disulphide bond C60–C10. Thus, tentatively we can identify E9–C10 and C60–K63–Y65 to be the sequences at the epitopic region.

Elimination of other Tyr and Lys residues can be done indirectly. Thus, K44 and K45 are eliminated because CNBr–hCG when treated with chymotrypsin retains almost 50% activity, a process wherein elimination of the peptide bond 33–47 is effected. K75 can be easily eliminated because of the stability of the epitope to chymotryptic digestion as well as its distance from a disulphide bridge. X-ray structure also puts this amino acid residue away from the core[49]. Lys 90 is eliminated on the basis of the stability of the epitope to CPase digestion of hCG. Lys 51, while near an acidic group is far away from a disulphide bond. These results point out to a region as probable epitope region for MAb 16 (Figure 5 *a*).

This conclusion is further substantiated by the observation that CPase digestion of chymotryptic digest causes 30% loss of activity. This is because the chymotryptic digest partially cleaves the Y65, exposing a tyrosine for CPase digestion, which causes the removal of Y65, resulting in loss of activity. Loss of activity to only 30% may indicate relative inaccessibility of Y65 to proteolysis, documented earlier[50]. Our own data (Table 1) also point to the inaccessibility of this region to proteolytic digestion.

Based on the above experimental data, we propose that the region corresponding to this epitope is as shown in Figure 5 *a*. The amino acids identified directly by chemical modification procedure are E9, C10, C60, K63 and Y65.

The data presented here convincingly show the importance of the disulphide bonds in the formation of this epitope. Hydrolysis of several peptide bonds does not alter the epitope structure, but reducing the disulphide bonds keeping each of the peptide bonds intact results in complete loss of activity and, hence, epitopic structure. This conformation-specific epitope seems to be almost a sequence-specific epitope through disulphide bridges (^{10}EC–C^{60}VAKSYN^{66}). It is notable that breaking the peptide bonds at 60–82 loop and/or 32–60 loop (by digestion with trypsin, chymotrypsin or CNBr) does not affect the integrity of this epitope. Thus, the integrity of this epitope is maintained more by disulphide bonds than by peptide bonds. The stabilizing force may not be the disulphide bond 10–60 exclusively. The rigid network of disulphide structure comprising cystine loops 7–31, 10–60, 59–87 and 84–32 (Figure 5) presents a structure (stable to proteolysis) which may be a strong force in keeping the integrity of this epitope. However, it is not possible to analyse the exact contribution of this network of the disulphide structures to the epitope stability in this case as the aminoacids 7–10 and 84–87 are not susceptible to any type of mild chemical or enzymatic cleavage, and selective reduction of these disulphide bonds has not been reported.

One of the arguments against the logic developed for analysis of epitopes by this method is that a modification at a site may induce conformational relaxation in the molecule at a distant epitope region. It is unlikely that a positive proof for or against this argument can be obtained. However, a large number of modifications like tryptic digestion (12 peptide bonds are cleaved), chymotryptic digestion (9 peptide bonds are cleaved) and TNBS modification (12 amino groups are modified) should induce sufficient long-range relaxations all over the molecule (all these modifications whether proteolysis or chemical modification are uniformly distributed, as can be seen by the sequence) and yet the effect on the epitopic activity is very specific. In addition, it is seen that modification of the amino group with different reagents like acetic anhydride, TNBS, biotin or Gld–DAE brings comparable loss of activity in spite of different side-chain groups (Table 3). These observations provide an indirect evidence that conformational relaxation at distant epitopic regions may not be the reason for the specific effects seen. This is possible considering that hCG has a very rigid, interwoven network of disulphide bridges. It would be interesting to know how the integrity of epitopes of a protein with few disulphide bridges is affected by modifications, as done above.

Currently available methods for identification of amino acid residues at conformation-specific epitopes rely on X-ray crystallographic data of Ag–Ab complex[51] or cross-species comparison using naturally occurring variants or mutant chimera[52,53]. These methods, though powerful, are adopted rather rarely as the method has no easy accessibility/adoptability to most of the

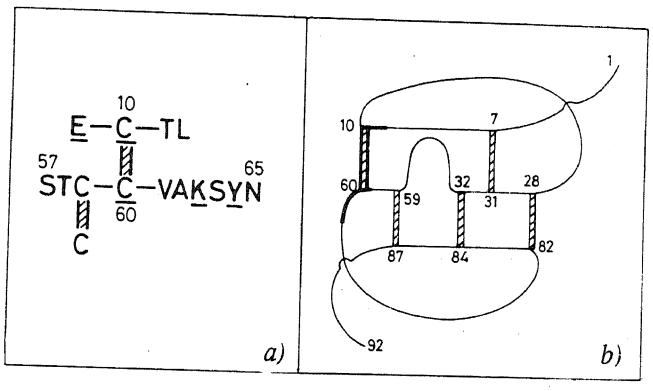


Figure 5. *a*, Epitope region of MAb 16. Underlined letters represent the amino acids identified by chemical/enzymatic modifications. Disulphides are presented by hatched bars. Numbers represent the position of amino acid in the primary sequence of the molecule. *b*, Schematic drawing of hCG α subunit highlighting the rigid disulphide core. Hatched bars represent the disulphide bonds. Epitope region 16 is indicated by the thick line (disulphide links are as per ref. 21 and 22).

laboratories. The method we have described above has an easy adoptability to routine laboratory with RIA facility, and is inexpensive and faster for identification of amino acid residues at epitopic regions. Our own experiments done with several epitopes of hCG indicate that this method is extendable to other epitopes of hCG and has the potential to become a general method for identification of epitopic regions for proteins with extensive disulphide bonds like gonadotropins. In addition, the method will have a greater reach to most of the laboratories when the RIAs are converted to EIAs, for which attempts are in progress.

1. Takahashi, H., Wilson, B., Ozturk, M., Motte, P., Strauss, W., Isselbacher, K. J. and Wands, J. R., *Cancer Res.*, 1989, **48**, 4573–4574.
2. Krichevsky, A., Birken, S., O'connor, J. F., Bikel, K., Schlatterer, J. P. and Canfield, R. E., *Endocrine*, 1994, **2**, 511–520.
3. Dirnnofer, S., Klieber, R., Leeuw, R. D., Bidart, J. M., Merz, W., Wick, G. and Berger, P., *FASEB J.*, 1993, **7**, 1381–1385.
4. Desmukh, U. S., Pal, R., Talwar, G. P. and Gupta, S. K. *J. Reprod. Immunol.*, 1993, **25**, 103–118.
5. Sela, M., *Science*, 1969, **166**, 1365–1374.
6. Berzofsky, J. A., *Science*, 1985, **229**, 932–940.
7. Benjamin, D. C., *Annu. Rev. Immunol.*, 1984, **2**, 67–101.
8. Barlow, D. J., Edwards, M. S. and Thornton, J. M., *Nature*, 1986, **322**, 747–748.
9. Attassi, M. Z., *Eur. J. Biochem.*, 1984, **145**, 1–20.
10. Attassi, M. Z., *Immunochemistry*, 1975, **12**, 423–428.
11. Muller, S., Plaue, S., Coupper, Z. and Van Regenmortel, M. H. V., *Mol. Immunol.*, 1986, **23**, 593–601.
12. Petterson, Y., *Mol. Biol. Reports*, 1992, **16**, 149–153.
13. Geyson, H. M., Rodda, S. J., Mason, T. J., Tribbick, G. and Schoofs, P. G., *J. Immunol. Methods*, 1987, **102**, 259–274.
14. Jung, G. and Beck-Sickinger, A. G., *Angew. Chem.*, 1992, **104**, 375–391.
15. Carter, J. M., Vanalbert, S., Lee, J., Lyon, J. and Deal, C., *Biotechnology*, 1992, **10**, 509–513.
16. Maeji, N. J., Bray, A. M. and Geyson, H. M., *J. Immunol. Methods*, 1990, **138**, 2213–2229.
17. Weiner, R. S. and Dias, J. A., *Mol. Cell Endocrinol.*, 1992, **85**, 41–52.
18. Schab, C., Twardek, A., Lo, T. P., Brayer, G. D. and Bosshard, H. F., *Protein Science*, 1990, **2**, 175–182.
19. Savoca, R., Schwab, C. and Bosshard, H. F., *J. Immunol. Methods*, 1991, **141**, 245–252.
20. Laver, W. G., Air, G. M., Webster, R. B. and Smith Gill, S. J., *Cell*, 1990, **61**, 553–556.
21. Reichlin, M., *J. Mol. Biol.*, 1972, **64**, 485–496.
22. Nisonof, A., Reichlin, M. and Margoliash, E., *J. Biol. Chem.*, 1970, **245**, 940–946.
23. Netther, H. J., Guldner, H. H., Szotstecki, C. and Will, H., *Scand. J. Immunol.*, 1990, **32**, 163–176.
24. St Clair, E. W., Kenan, D., Burch, J. A., Keene, J. D., Pisetsky, D. S., *J. Immunol.*, 1990, **144**, 3868–3876.
25. Jemmerson, R. and Blankenfield, R., *Mol. Immunol.*, 1989, **26**, 301–307.
26. Van Regenmortel, M. H. V., *Phil. Trans. R. Soc. Lond. B*, 1989, **323**, 451–466.
27. Blundell, T. L., Sibanda, B. L., Sternberg, M. J. E and Thornton, J. M., *Nature*, 1987, **326**, 347–352.
28. Amit, A. G., Meriuzza, R. A., Philips, S. E. V. and Poljak, R. J., *Science*, 1986, **233**, 747–753.
29. Sheriff, S., Silverton, E. W., Padlon, E. A., Cohen, G. H., Smith Gill, S. J., Finzel, B. C. and Davis, D. R., *Proc. Natl. Acad. Sci. USA.*, 1987, **84**, 8075–8079.
30. Tulip, W. R., Varghese, J. M., Webster, R. G., Air, G. N., Laver, W. G. and Colman, P. M., *Cold Spring Harbor Symp. Quant. Biol.*, 1990, **54**, 257–263.
31. Padlan, E. A., Silverton, E. W., Sheriff, S., Cohen, G. H., Smith Gill, S. J. and Davis, D. R., *Proc. Natl. Acad. Sci. USA*, 1989, **86**, 5938–5942.
32. Wilson, J. E., *Methods Biochem. Analysis*, 1991, **35**, 207–248.
33. Horsfall, A. C., Hay, F. C., Soltys, A. J. and Jones, M. G., *Immunol. Today*, 1991, **12**, 211–213.
34. Fraker, P. J. and Speck, J. C., *Biochem. Biophys. Res. Commun*, 1978, **80**, 849–857.
35. Dighe, R. R., Murthy, G. S. and Moudgal, N. R., *J. Immunol. Methods*, 1990, **122**, 229–236.
36. Murthy, G. S., Lakshmi, B. S. and Moudgal, N. R., *J. Biosci.*, 1989, **14**, 9–20.
37. Riordan, J. F. and Vallee, B. L., *Methods Enzymol.*, 1972, **25**, 494–499.
38. Habeeb, A. F. S. A., *Anal. Biochem.*, 1966, **14**, 328–336.
39. Beyer, E. A. and Wilcheck, M., *Methods Enzymol.*, 1990, **184**, 138–160.
40. Riordan, J. F. and Vallee, *Methods Enzymol.*, 1972, **25**, 515–521.
41. Spade, T. F., Witkop, B., Degani, Y. and Pachornik, A., *Adv. Protein Chem.*, 1970, **24**, 97–260.
42. Rice, R. H. and Etzler, M. E., *Biochemistry*, 1975, **14**, 4093–4098.
43. Sashidar, R. B., Capoor, A. K. and Ramana, D., *J. Immunol. Methods*, 1994, **167**, 121–127.
44. Garnier, J., Salesse, R. and Pernollet, C. J., *FEBS Lett.*, 1974, **45**, 166–171.
45. Mori, K. F. and Hollands, T. R., *J. Biol. Chem.*, 1971, **246**, 7223–7229.
46. Wilson, D. E., Sun, S., Ozturk, M. and Wands, J. R., *J. Immunol. Methods*, 1991, **139**, 55–64.
47. Underwood, P. A., *J. Immunol. Methods*, 1985, **85**, 295–323.
48. Steward, M. W. and Lew, A. M., *J. Immunol. Methods*, 1985, **78**, 173–190.
49. Lapthorn, A. J., Harris, D. C., Little John, A., Lustbader, J. W., Canfield, R. E., Machin K. J., Morgon, F. J. and Issacs, N. W., *Nature*, 1994, **369**, 455–461.
50. Birken, S., Kolks, M. A. G., Amr, S., Nisula, B. and Puet, D., *J. Biol. Chem.*, 1986, **261**, 10719–10727.
51. Tulip, W. R., Varghese, J. N., Webster, B. G. and Colman, P. M., *J. Mol. Biol.*, 1992, **227**, 149–159.
52. Bidart, J. M., Troalen, F., Bousfield, G. R., Bouhon, C. and Bellet, D., *Endocrinology*, 1989, **124**, 923–929.
53. Moyle, W. R., Matzuk, M. M., Campbell, R. K., Cogliani, E., Dean-Emig, D. M., Krichevsky, A., Barnett, R. W. and Boime, I., *J. Biol. Chem.*, 1990, **265**, 8511–8518.

ACKNOWLEDGEMENTS. We thank Dr Rajan Dighe for providing MAb culture supernatants, Prof. N. R. Moudgal for his encouragement, CSIR for Research Fellowship to N. V. and J. Ramesh and J. K. Tiwari for their help in the analysis of CD spectra. The financial assistance of DBT and Malladi Research Foundation is gratefully acknowledged.

Received 24 March 1995; revised accepted 5 May 1995

On minimum metallic diffusivity in disordered conductors

N. Kumar and A. M. Jayannavar*

Raman Research Institute, Bangalore 560 080, India
*Institute of Physics, Bhubaneswar 751 005, India

It is shown that the formal consistency of a locally diffusive motion on the metallic side of the mobility edge for a disordered system with the Schrödinger equation implies a minimum metallic diffusivity $D_{min} = \hbar/2m$. Its implification for the physics near the mobility edge is discussed.

METALLIC conductivity σ implies diffusivity D through the Einstein relation for the degenerate electron system, namely, $\sigma = e^2 D N_F$. For a quench-disordered conductor, this poses the question of formal consistency of the diffusive equation for the probability density with the Schrödinger equation for the probability amplitude. In what follows we address this question and find the surprising result that a necessary condition is $D_{min} = \hbar/2m$.

Consider a phase space wave packet initially localized around the origin and constituted from the extended energy eigenstates close to the Fermi energy (counted as zero of energy) which is taken to lie just on the metallic side of the mobility edge for the quench-disordered system.

Rewrite the Schrödinger equation for the wave packet

$$\psi(x, t) = R(x, t)\, e^{i\phi(x,t)} \equiv e^{\alpha(x,t) + i\beta(x,t)} \tag{1}$$

as

$$-\hbar \frac{\partial \phi}{\partial t} = \frac{\hbar^2}{2m} (\nabla \phi)^2 + V - \frac{\hbar^2}{2m} \frac{\nabla^2 R}{R} \tag{2}$$

along with the continuity equation

$$\frac{\partial \rho}{\partial t} + \frac{\hbar}{m} \nabla \cdot (\rho \nabla \phi) = 0 \tag{3}$$

with $\rho = \psi^* \psi$.

Now, we would like the continuity equation for the probability density to converge asymptotically to the diffusive form

$$\frac{\partial \rho}{\partial t} = D \nabla^2 \rho. \tag{4}$$

Comparing eqs. (3) and (4), we at once get a condition for this to be so in terms of α and β, which is

$$\frac{\hbar}{2m} \nabla \beta = D \nabla \alpha \tag{5}$$

assuming $\nabla\alpha$ to be irrotational.

We now substitute in eq. (2) from eq. (5) and rearrange to get

$$-\hbar \frac{\partial \beta}{\partial t} = V - \frac{\hbar^3}{4m^2 D} \nabla^2 \beta + \frac{\hbar^2}{2m} \left(1 - \left(\frac{\hbar}{2mD}\right)^2\right) (\nabla \beta)^2. \tag{6}$$

Now, averaging over the disorder V and noting that odd-symmetric quantities, i.e. $\langle \beta \rangle = \nabla^2 \langle \beta \rangle = \langle V \rangle = 0$, while $\langle (\nabla \beta)^2 \rangle \neq 0$, we necessarily have

$$D = \left(\frac{\hbar}{2m}\right). \tag{7}$$

Here, of course, we have regarded D as a self-averaging quantity. This is reasonable as we know that on the metallic side of the mobility edge in three dimensions the conductance fluctuations are universal ($\sim e^2/h$), but the conductivity fluctuation is $\sim (e^2/h)(1/L)$ and, therefore, vanishes in the infinite sample size (L) limit. We must note that our sample size L is always taken to be essentially infinite ($\gg$ any correlation length ξ that may exist as we approach the mobility edge). This ensures that we are always in the local diffusive regime in that the mean-squared displacement grows linearly in time as $2Dt$. For $L \leq \xi$, the diffusion constant itself can, of course, be L-dependent. This is, however, quite different from the anomalous diffusion, where the mean-squared displacement does *not* grow linearly with time. Further, we have insisted on the local diffusion equation (4) and this forces us physically to the limit when the mean free path equals the microscopic length – the Fermi wavelength ~ the lattice discreteness. The latter implies that we should be closest to the mobility edge, but on the metallic side. Thus, $\hbar/2m$ is to be taken as the minimum possible value of D to ensure local diffusion.

As to the general conceptual question of how diffusion of an initially localized wave packet may come out of the Schrödinger equation with quenched (static) disorder, we consider the following. Of course, in the static random potential there is no dissipation and hence no irreversibility that one normally associates with the inelastic (incoherent) scatterings. Indeed, in the strictly fine-grained sense there is no irreversible diffusive spread of the electron wave packet, and eq. (3) is time-reversal-symmetric (i.e. under $t \to -t$ and $\phi \to -\phi$). This old problem[1] is the same as that discussed recently by Imry[2] in the context of the nonzero real part of the dc conductivity. The point is that for the extended states the level spacing scales to zero as the system size tends to infinity, i.e. the poles produce a cut. Then arbitrarily small level broadening will ensure the phase-mixedupness of the otherwise pure state and thus ensure irreversibility in the long-time limit. This is really an analogue of the collisonless Landau damping! Indeed, Mott and Kaveh[3] have discussed the diffusion equation as arising from the Schrödinger equation for a disordered system.

The above derivation strongly suggests that there exists a minimum metallic diffusivity D_{min} having a universal value $\hbar/2m$ *whenever* a mobility edge is known to exist. The question now is whether or not this implies a minimum metallic conductivity[4] σ_{min} through the Einstein relation $\sigma = e^2 D N_F$. This brings us to an oft-stated but seldom proven theorem[5], namely, that the localized states cannot coexist with (be at the same energy) as the extended states. And this notwithstanding the well-known counterexamples of the Wigner–von Neumann construction of such states[6]. The latter seem to be dismissed as nongeneric. If the coexistence is denied, the N_F on the metallic side of the mobility edge counts the total density of states (there being no coexisting localized states) and then $D_{min} = \hbar/2m$ would imply a $\sigma^{3D}_{min} = e^2 D_{min} N_F$. In this connection it is worth noting that the ultrathin films of Bi on Ge substrate show[7] a limiting normal metallic conductivity of $4e^2/h$ (corresponding to a sheet resistance $\sim 6.5\ k\Omega/\square$) for the separatrix dividing the insulating and the superconducting phases at 0 K. This may well be associated with the preexisting local electron pairs where the correlated motion of the pairing electron escapes the tyranny of two-dimensional (2D) localization. Now, it is interesting that our universal $D_{min} = \hbar/2m$, along with $N_F = (m/\pi\hbar^2)$ for $d = 2$ does indeed give $\sigma^{2D}_{\square} = 4e^2/h$. We would like to call it the σ^{2D}_{min} of Mott recovered through delocalization by pairing in two dimensions. This may be the case for all insulator–superconductor transitions with a normal-state separatrix for the ultrathin films.

There is, however, the question of the fine structure of the mobility edge for $d = 3$, where the conductivity is experimentally believed to vanish continuously[8] at the mobility edge in a narrow critical region with a critical exponent s. How is this consistent with $D_{min} \neq 0$? A possible solution consistent with $D_{min} \neq 0$ is to assume that $N_F = N_F(\text{localized}) + N_F(\text{extended})$. What then enters the Einstein relation is, of course, the $N_F(\text{extended})$, which can vanish at the mobility edge keeping the total N_F continuous and noncritical, as is known to be the case. It is not clear how this two-spectral fluid hypothesis[5], i.e. $N_F(E) = N_F(\text{localized}) + N_F(\text{extended})$, really differs from the asymptotic form of the resonance-like wave function in three dimensions adduced by Kaveh and Mott[9]. Such quasilocalized resonant states are strongly suggested by some recent works[10,11]. We think that a proper discussion on this point is yet to take place.

Finally, two technical remarks are in order. First, the validity of our eqn. (5) at any point x demands that x not be a node where $\psi(x)$ vanishes. But this condition is naturally obeyed by all diffusive states. Indeed, vanishing of $\psi(x)$ would imply vanishing of $\rho(x) = \psi^*(x)\,\psi(x)$, which is ruled out by the fact that $\rho(x)$ is requested to obey the diffusion equation (3). The second point concerns the absolute value $4e^2/h$ for the sheet conductance of films referred to in the text, i.e. for Bi films on Ge. The point is that the effective band mass does not appear in the expression, and the correction to the density of states due to disorder or interaction is cancelled by the vertex correction, at least perturbatively – the well-known cancellation theorem valid in the absence of a pseudo gap[12].

1. Thirring, W., in *Preludes in Theoretical Physics* (De-Shalit, A., Feshbach, H. and van Hove, L., eds), North-Holland, Amsterdam, 1966, p. 266.
2. Imry, Yoseph, in *Directions in Condensed Matter Physics* (Grinstein, G. and Mazenko, G., eds), World Scientific, Singapore, 1986, p. 101.
3. Mott, N. F. and Kaveh, M., *Adv. Phys.*, 1985, **34**, 329.
4. Mott, N. F., *Phil. Mag.*, 1972, **26**, 1015.
5. Phillips, J. C., *Phil. Mag.*, 1983, **B47**, 407.
6. Ballentine, L. E., *Quantum Mechanics*, Prentice-Hall, London, 1990, p. 205.
7. Haviland, D. B., Liu, Y. and Goldman, A. M., *Phys. Rev. Lett.*, 1989, **62**, 2180.
8. Lee, P. A. and Ramakrishnan, T. V., *Rev. Mod. Phys.*, 1985, **57**, 287.
9. Kaveh, M. and Mott, N. F., *Phil. Mag.*, 1983, **47**, 577.
10. Karpov, V. G., *Phys. Rev.*, 1993, **B48**, 4325.
11. Kumar, N., in *Ordering Disorder* (Srivastava, V., Bhatnagar, A. K. and Naugle, D. G., eds), AIP Conf. Proc., 286, New York, 1994, p. 3.
12. Mott, N. F., *Metal–Insulator Transitions*, Taylor and Francis, London, 1974, p. 28.

Received 18 February 1995; revised accepted 19 May 1995

An *ab initio* topographical investigation on the molecular electrostatic potential of some chemical mutagens

Apurba K. Bhattacharjee, Savita S. Pundlik* and Shridhar R. Gadre*

Department of Chemistry, Lady Keane College, Shillong 793 001, India
*Department of Chemistry, University of Poona, Pune 411 007, India

A detailed topographical investigation on the molecular electrostatic potentials (MESPs) of different conformers of acetaldehyde, nitrous acid and hydroxylamine has been carried out at the *ab initio* SCF level using TZ2p, 6-31G* and STO-3G basis sets. In general, large regions of negative potential have been observed. An attempt has been made to correlate these potentials with biological activities of the molecules. Mutagenic and toxicological properties appear to be related to the presence of these large negative zones.

MOLECULAR electrostatic potential (MESP) is now an established tool for the study of the stereoelectronic

attributes and properties of binding molecules which would significantly contribute to the elucidation of biochemical phenomena[1–3]. Even long-range electrostatic interactions have now been shown to be important for a number of biological reactions including folding of proteins, determination of pKa's of the ionizable residues etc.[4]. The affinity of a particular molecule for a specific receptor is shown to depend on the MESP characteristics[5]. This ability of a molecule for interaction with the receptor is essentially known as the 'recognition' process, in which the receptor 'recognizes' the characteristic features of the molecule in order to promote their interaction. MESP can, therefore, be used for predicting structure–activity relationships when a receptor-mediated response is supposed to give rise to toxicological, mutagenic or carcinogenic effects even without a detailed knowledge of the receptor[6]. MESP provides the exact value for the electrostatic component of the interaction energy between an unperturbed charge distribution and a proton[7]. Moreover, MESP gives a general picture of the reactive properties of the molecule with a reasonable computational cost and this explains its large number of applications in very different fields of physics, chemistry, biochemistry and pharmacology[8]. In a recent report Zhu *et al.*[9] have shown that radiosensitizers can be made based on MESP data, thereby indicating its possible applications in the study of molecular toxicology. Efficient radiosensitizers are characterized by a wide and deep negative potential in their MESP profiles, while this potential is narrow for compounds devoid of significant activity[9].

The MESP $V(r)$ at a point r due to nuclear charges $\{Z_A\}$ at $\{R_A\}$ and the electronic density $\rho(r)$ of the molecule is given by

$$V(r) = \sum_A \frac{Z_A}{|R_A - r|} - \int \frac{\rho(r')}{|r - r'|}\, d^3r'. \qquad (1)$$

The first term on the right-hand side of this equation represents the bare nuclear potential while the second one represents the electronic density function of the molecule. The sign of $V(r)$ in any particular region depends on whether the nuclear or the electronic effects are dominant there.

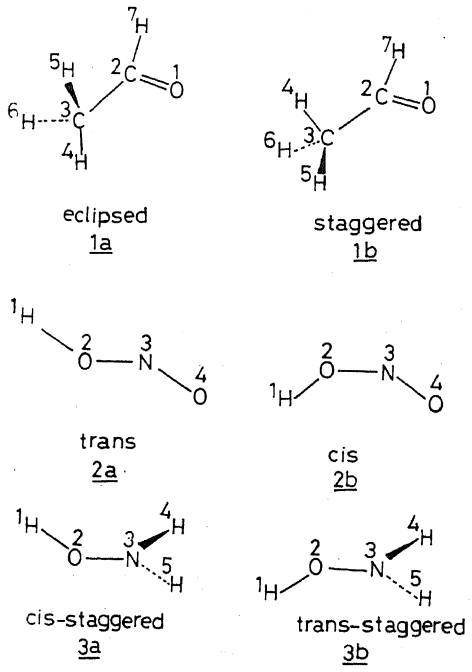
The topographical analysis of $V(r)$ involves identification and location of the critical points (CPs)[10,11], the points at which $\nabla V(r) = 0$. A CP in the MESP topography specifies a point at which the internal electric field (molecular electric field or all the partial derivatives of MESP) is exactly zero. This implies that the numerical value of the molecular electric field can be used for locating CPs in the MESP topography. The rank of the CP is given by the number of nonzero eigenvalues of the Hessian matrix $\boldsymbol{A}$, whose elements are given by

$$A_{ij} = \frac{\partial^2 V}{\partial r_i\, \partial r_j}\Bigg|_{r = r_{CP}}. \qquad (2)$$

A CP of rank 3 (known as a nondegenerate CP) is characterized by the algebraic sum of the signs of eigenvalues. Thus, there are four possible types of CPs: (3, –3), (3, –1), (3, +1) and (3, +3). The (3, –3) CP cannot exist in $V(r)$ except at the nuclear locations[11]. A (3, –1) or (3, +1) CP would correspond to saddle points whereas a (3, +3) CP is due to the local minimum of $V(r)$ beyond the van der Waals surface of the molecule.

The MESP minima can give an idea of the deprotonation sites as well as the interaction possibilities at the electrophilic centres of the receptor. Since MESP reflects the molecular interaction energy with a test positive point charge, the (3, +3) CPs would correspond to their proton affinities[12]. This would imply a threshold interaction energy of the atoms in the molecule to extract a proton from the proton donor site on the receptor. Good correlations between proton affinity or hydrogen bond acceptor parameters and MESP have been established in the recent past[13].

Chemical mutagens act by modifying the bases of DNA. Nitrous acid, hydroxylamine and acetaldehyde are all well-known mutagens[14]. Nitrous acid causes AT ↔ GC transitions whereas hydroxylamine is a highly specific mutagen which reacts almost exclusively with cytosine to give a derivative that pairs with adenine rather than with guanine and so it produces a unidirectional transition of CG to AT[14]. Acetaldehyde has also been reported[15] to be highly toxic, comparable to nitrous acid and hydroxylamine. These molecules, apart from their toxicities, are also known to be efficient proton removers and deaminating agents[16]. Further, these molecules are also interesting because there is a strong evidence that they can act as mutagens most likely at the same site[12].



The objective of the present study is to assess the importance of three-dimensional MESP topographical analysis of acetaldehyde (1a, 1b), nitrous acid (2a, 2b) and hydroxylamine (3a, 3b) at a higher level of theory in the mutagen–receptor interaction process. Although electron density maps of some of these molecules[17] have been reported earlier, to our knowledge detailed topographical analysis at this level of theory has not been reported so far in the literature.

We have used *ab initio* optimized geometries of three molecules in their different conformations[17–19] and obtained the wave functions (with TZ2p, 6-31G* and STO-3G basis sets) by using the INDMOL package[20] on a 64-node transputer-based machine called PARAM[21]. The MESP calculations were carried out using these wave functions on the same computer with the help of a program developed by Shirsat *et al.*[22]. The computed electrostatic potentials are at a distance of about 3 a.u. surrounding the molecule, which corresponds approximately to the van der Waals radius of the largest atom in this group of molecules (oxygen)[23]. The potential at this surface would usually be encountered by the receptor during the 'recognition' process. The basis set dependence of the MESP was checked by using different basis sets in these calculations. The CPs were located using the analytical expressions for the gradient of $V(r)$ given by Gadre *et al.*[10,24].

In order to have a quantitative measure of the negative region typical for biological recognition interaction[3], we have calculated the volume inside a closed surface having an MESP energy of –15.6 kcal/mol. This is done by calculating MESP values on a 64-node PARAM machine[21] inside a rectangular parallelopiped enclosing an isosurface of this energy and sampling the points through a random search (Monte Carlo method). The volumes thus obtained for all the molecules studied are indicated in Table 1.

In Table 1, the CP characteristics of the molecules studied at TZ2p level are presented. The calculated MESP minima (most negative values) are beyond the van der Waals surface of the molecules. The electrostatic potential maps of the molecules are presented in Figure 1.

In acetaldehyde, practically no energy difference is found between the MESP critical points of the eclipsed and the staggered conformations, which indicates that the MESP profile is independent of methyl rotation, although methyl barrier to internal rotation around the C=O bond has been calculated at the 6-31G* level to be 1.06 kcal/mol (ref. 17). The two negative-valued minima are found in the lone pair regions of oxygen in both the rotamers. The TZ2p MESP maps of both the rotamers show the presence of strong negative potential zones laterally around oxygen encompassing a broad region which might be linked to the strong biological activity of the molecule. The negative regions

Table 1. Types and locations of negative-valued MESP critical points (CPs) calculated at TZ2p level

Molecule*	CP type	Coordinates (a.u.) x	y	z	MESP value (kcal/mol)	Volume (a.u.)^3
Acetaldehyde						
eclipsed[a]	(3, +3)	–1.42	–1.82	0.0	–52.8	
	(3, +1)	–2.43	0.0	0.0	–46.5	239.3
	(3, +3)	–1.47	1.77	0.0	–51.4	
staggered[b]	(3, +3)	–1.42	–1.82	0.0	–53.0	
	(3, +1)	–2.43	0.0	0.0	–46.8	241.4
	(3, +3)	–1.47	1.77	0.0	–51.6	
Nitrous acid						
trans-[c]	(3, +3)	1.20	4.40	0.0	–21.9	
	(3, +1)	2.87	5.78	0.0	–11.9	
	(3, +3)	5.81	4.63	0.0	–36.9	
	(3, +1)	7.54	2.72	0.0	–25.9	156.0
	(3, +3)	6.25	0.47	0.0	–36.5	
	(3, +1)	4.82	–1.20	0.0	–22.9	
	(3, +3)	3.40	–1.92	0.0	–28.5	
cis-[d]	(3, +3)	–0.56	4.60	0.0	–18.4	
	3, +3)	3.04	6.06	0.0	–30.2	
	(3,+1)	5.72	4.65	0.0	–10.3	34.0
	(3, +3)	5.59	1.96	0.0	–24.4	
	(3, +1)	5.56	–0.93	0.0	–7.5	
	(3, +3)	3.18	–2.06	0.0	–20.7	
Hydroxylamine						
cis-staggered[e]	(3, +3)	4.92	2.44	0.0	–68.7	
	(3, +1)	4.80	–0.18	0.0	–38.0	
	(3, +1)	3.32	–1.77	0.0	–53.1	278.8
	(3, +3)	3.19	–1.64	–0.8	–53.3	
	(3, +3)	3.19	–1.64	0.8	–53.3	
trans-staggered[f]	(3, +3)	3.03	–1.91	0.0	–51.2	
	(3, +3)	1.20	4.00	0.0	–63.0	123.0

*Nuclear coordinates (in a.u.) of some of the atoms are:
[a]O (0,0,0), C (1.18,0,0), C (2.03,1.24,0), H (1.40,2.12,0).
[b]O (0,0,0), C (1.18,0,0). C(2.03,1.25,0), H (1.82,1.85,0.87).
[c]H (0,0,0), O (0.95,0,0), N (1.40,1.23,0), O (2.56,1.29,0).
[d]H (0,0,0), O (0.96,0,0), N (1.59,1.12,0), O (0.9,2.09,0).
[e]H (0,0,0), O (0.95,0,0), N (1.36, 1.26,0), H (0.99,1.71,–0.84).
[f]H (0,0,0), O (0.95,0,0), N (1.47,1.22,0), H (2.05,1.38,0.83).

associated with the carbonyl oxygen extend to a fairly large region; the minima representing the lone pairs have a value around –51.0 kcal/mol separated by a saddle corresponding to about –46.0 kcal/mol.

In the case of *trans*-nitrous acid, seven distinct CPs with a deepest minimum of about –37.0 kcal/mol are observed. A large extended area of about –20.0 kcal/mol is observed which is most likely to be the characteristic feature of the 'recognition' interaction. For *cis*-nitrous acid, six CPs were located, with the deepest minimum at –30.0 kcal/mol. A broad extended area of about –8.0 kcal/mol is observed as the saddle energy.

Cis-staggered hydroxylamine shows three minima, two of which are at –68.0 and the other one is at –53.0 kcal/mol. A saddle energy range of about –40.0 to –50.0 kcal/mol is found to be extended over a large area on only one side of the molecule. In the case of *trans*-staggered hydroxylamine, however, the MESP

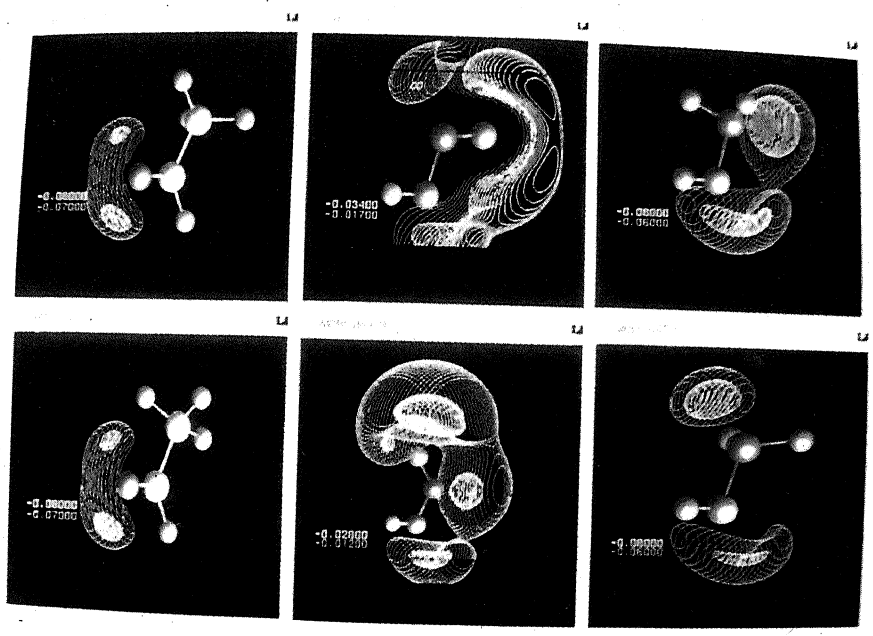


Figure 1. The molecular electrostatic potential contours of: *a*, acetaldehyde eclipsed (upper) and staggered (lower). The green colour corresponds to a value of –44.0 kcal/mol, while the pink one corresponds to –50.2 kcal/mol. *b*, *Trans*-nitrous acid (upper) showing green colour contours of –10.6 kcal/mol and pink colour contours of –21.3 kcal/mol, and *cis*-nitrous acid (lower) showing green colour contours of –7.5 kcal/mol and pink colour contours of –17.5 kcal/mol. *c*, Hydroxylamine *cis*-staggered (upper) and *trans*-staggered (lower). The green colour corresponds to a value of –37.6 kcal/mol, while the pink one corresponds to –50.2 kcal/mol.

minima are localized on both sides of the molecule, one near oxygen and the other near nitrogen, with the potential values of –50.8 and –62.7 kcal/mol, respectively. The saddle energy of about –40.0 kcal/mol is found to be well extended on both sides of this molecule.

Calculations made at the 6-31G* level could reproduce similar MESP profiles for all the molecules. For *cis*-staggered hydroxylamine, however, the two minima off the *xy* plane obtained with TZ2p basis could not be distinguished and they appeared to have converged into a single one in the plane.

Although STO-3G calculations show qualitative similarity of the MESP features of the molecules, the locations and the potential values of the CPs differ significantly. In acetaldehyde, for example, the deepest minimum has a higher value of MESP (–34.0 kcal/mol) and in *cis*-nitrous acid, it is as low as – 49.0 kcal/mol. Thus, the choice of STO-3G basis sets for MESP calculations of such molecules should perhaps be made with caution.

The magnitudes of the negative-valued minima of all the molecules may be correlated with their proton affinities. This is because MESP reflects the molecular interaction energy with a positive point charge; therefore, this MESP minimum threshold would correspond to the proton affinity of the molecules (the amount of energy necessary to extract a proton from the proton donor site on the receptor). Such correlations have been reported earlier for a series of oxygen- and nitrogen-containing molecules[12,25].

In order to compare the biological activities, we have scanned volumes enclosed within the negative region having potentials in the neighbourhood of –15.0 kcal/mol for all the molecules. A value of such a magnitude has been reported by Murray *et al.*[3] as optimum potential minimum to explain the toxicities of several dibenzo-*p*-dioxins. In our scan, we too observe (Table 1) a very large area being encompassed with this energy. It may thus be reasonable to infer that both the depth of the MESP minima and the magnitude of the 'volumes' together play a crucial role in the biological interaction process. From an inspection of the 'volume enclosure' of –15.6 kcal/mol, it is apparent that both the rotamers of acetaldehyde (eclipsed 239.8 a.u^3 and staggered 241.4 a.u^3) and *cis*-staggered hydroxylamine (278.8 a.u^3) are more prone to biological recognition interaction than *trans*-staggered hydroxylamine (123.0 a.u^3), *trans*-nitrous acid (156.0 a.u^3) and *cis*-nitrous acid (34.0 a.u^3). It may be summarized from the above results that *cis*-staggered hydroxylamine would be biologically more active, followed by the two acetaldehyde rotamers. Apart from having high negative potential regions the molecules are much smaller in size, which enable them to a greater accessibility to the receptor site. Thus, three factors – (i) the magnitude of the deepest negative potential, (ii) a large optimum negative potential zone and (iii) a

smaller size (consequently, a better steric accessibility) – appear to be crucial for the molecules to form an energetically favourable complex with the receptor which would interfere with the subsequent reactions taking place in the receptor.

The computational results of this theoretical MESP analysis on the different conformers of acetaldehyde, nitrous acid and hydroxylamine indicate that three factors – (i) the magnitude of the deepest negative potential, (ii) a large 'optimum' negative potential zone and (iii) small size – would contribute largely towards their biological activity. The MESP analysis may thus be used as a tool to determine the degree of mutagenic capacity *vis-à-vis* deprotonating as well as deaminating abilities of these molecular mutagens.

Although electrostatic potential patterns shown in this study can offer a rich informative description about the capacity of the molecules to generate stereoelectrostatic forces, they fail to describe their ability to generate hydrophobic bonds and dispersion interactions. Therefore, the MESP profiles in the present investigation would reveal only the stereoelectronic components of intermolecular recognition forces. Thus, further sophisticated calculations including solvent effects are required to prove or disprove the proposed reasons for the biological potency of the molecules.

1. Guha, S., Majumdar, D. and Bhattacharjee, A. K., *J. Mol. Struct. (Theochem.)*, 1992, **88**, 61–74.
2. Majumdar, D., Bhattacharjee, A. K., Das, K. K. and Guha, S., *J. Mol. Struct. (Theochem.)*, 1993, **288**, 41–53.
3. Murray, J. S., Evans, P. and Politzer, P., *Int. J. Quant. Chem.*, 1990, **37**, 271–275.
4. Nathrup, S., Boles, J. and Reynolds, J., *Science*, 1988, **241**, 67.
5. Thompson, C. and Brandt, R., *Int. J. Quant. Chem., Quant. Biol. Symp.*, 1983, **10**, 357–361.
6. Richard, A. M., *J. Comput. Chem.*, 1991, **12**, 959–962.
7. Alhanbra, C., Luque, F. J. and Orozco, M., *J. Phys. Chem.*, 1995, **99**, 3084–3087.
8. Politzer, P. and Murray, J. S., in *Reviews in Computational Chemistry* (eds Lipkowitz, K. B. and Boyd, D. B.), VCH Publishers, New York, 1991, vol. 2.
9. Zhu, A., Xu, S., Huang, J. and Luo, Z., *Int. J. Radiat. Biol.*, 1989, **56**, 893–895.
10. Gadre, S. R., Kulkarni, S. A. and Shrivastava, I. H., *J. Chem. Phys.*, 1992, **96**, 5253–5256.
11. Pathak, R. K. and Gadre, S. R., *J. Chem. Phys.*, 1990, **93**, 1770–1772.
12. Luque, F. J., Sanz, F., Illas, F., Ponplana, R. and Smeyers, Y. G., *Eur. J. Med. Chem.*, 1988, **23**, 7–10.
13. Murray, J. S., Ranganathan, S. and Politzer, P., *J. Org. Chem.*, 1991, **56**, 3734–3738.
14. Stryer, L., *Biochemistry*, 3rd edn, Freeman & Co., New York, 1988, p. 677.
15. Robbins, S. L., Cotran R. S. and Kumar, V., *Pathological Basis of Diseases*, 3rd edn, Saunders, London, 1984, pp. 122, 922.
16. Potapov, V. M. and Tartarinchic, S. N., *Organic Chemistry* (translation), Mir, Moscow, 1979, p. 222.
17. Wiberg, K. and Martin, E., *J. Am. Chem. Soc.*, 1985, **107**, 5035–5040.
18. Gordon, M. S., *J. Am. Chem. Soc.*, 1969, **91**, 3122–3125.
19. Cox, A. P. and Kuczkowski, L., *J. Am. Chem. Soc.*, 1966, **88**, 5071–5075.
20. Shirsat, R. N., Limaye, A. C. and Gadre, S. R., *J. Comput. Chem.*, 1993, **14**, 445–449.
21. Eknath, P. R., Bhasin, L. and Degwekar, A., in *Advanced Computing* (ed. Bhatkar, V. P.), Tata McGraw Hill, New Delhi, 1991, p. 86.
22. Shirsat, R. N., Bapat, S. V. and Gadre, S. R., *Chem. Phys. Lett.*, 1992, **200**, 373–375.
23. Bondi, A., *J. Phys. Chem.*, 1964, **68**, 441–445.
24. Gadre, S. R. and Shrivastava, I. H., *J. Chem. Phys.*, 1991, **94**, 4384–4390.
25. Kollman, P. and Rothenberg, S., *J. Am. Chem. Soc.*, 1977, **99**, 1333–1336.

ACKNOWLEDGEMENTS. One of us (AKB) thanks the Department of Science and Technology (DST), New Delhi, for a visiting SERC fellowship to carry out the work. The financial assistance from the Council of Scientific and Industrial Research (CSIR) is gratefully acknowledged. We are extremely thankful to the Centre for Development of Advanced Computing (C-DAC), Pune, for providing the computational facilities.

Received 19 September 1994; revised accepted 18 May 1995

Record of 1.66-Ga-old marine oxido-reducing conditions from an unusual Sargipali Pb–Cu deposit (India): A preliminary appraisal through petrogenesis

Rajesh K. Vishwakarma

Department of Geology, Banaras Hindu University, Varanasi 221 005, India

In the Proterozoic marine reducing environment the evidence of oxidizing condition from petrogenesis of some unusual metamorphic rocks like calcite–brucite, quartz–diopside and diopside-rich skarn has not been reported. The results of this contribution imply that, despite complete alteration of their precursor sedimentary rock composition, there is little change in the original behaviour of the rare-earth elements. Primarily on this basis, information about the changing Proterozoic oxidation state in the reducing seawater environment is put forth on the basis of Ce anomaly. The observation may be significant in view of the worldwide rarity of data concerning oxido-reduction characteristics of the Proterozoic marine realm. In addition, a new documentary evidence for $\delta^{18}O$ and $\delta^{13}C$ isotopic depletion in the marble developed due to metamorphism of the pure dolomite rock under at least 600°C at 6 kb *PT*-conditions is also brought out. Until now it was generally believed[1] that the pure carbonate rocks do not show a net lowering of these isotopes under metamorphic conditions.

THE 1.66-Ga-old stratiform Sargipali sulphide deposit[2] of the Gangpur Group, eastern India (Figure 1), is a

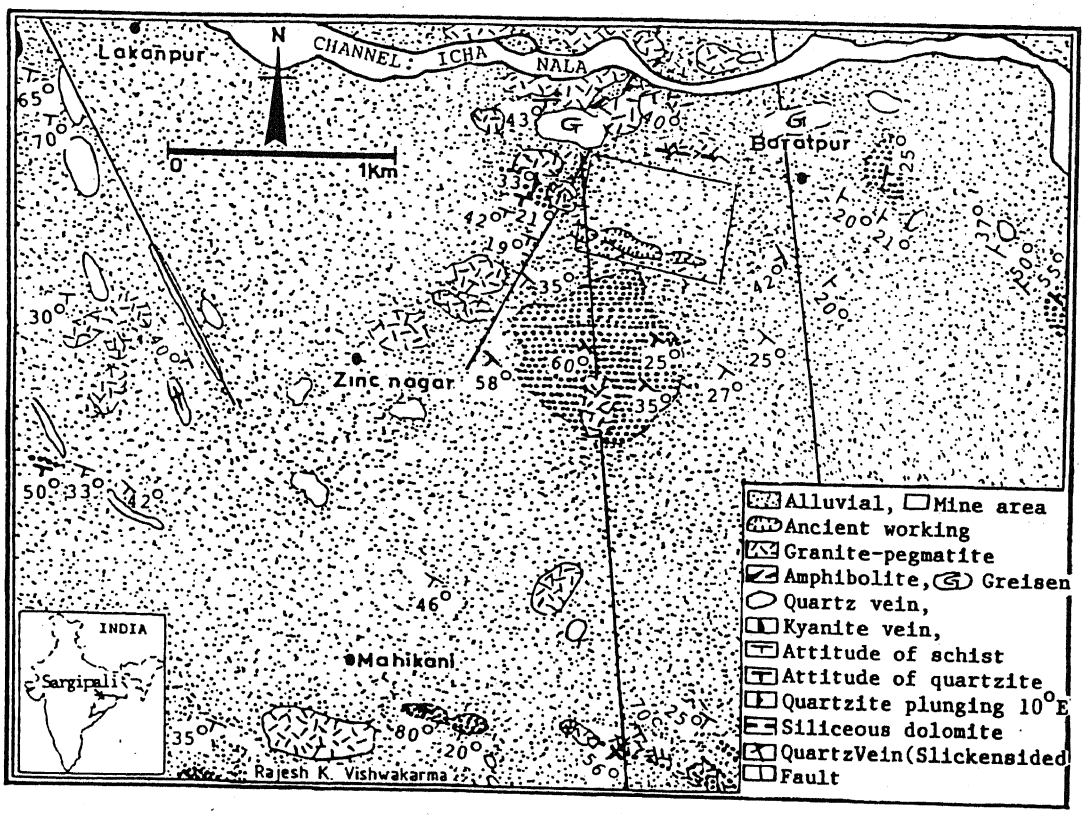


Figure 1. Geological map of the Sargipali Pb–Cu mine area. Inset shows the location of the study area.

syngenetic deposit which is formed due to sedimentary exhalative activity in response to the slow-spreading rift system[3]. Palaeoenvironmental study indicates that the area represents intertidal to supratidal marine flat environment[3]. This interpretation is based on polymictic metaconglomerate and current bedded structures (within meta-arenaceous lithounit), intercalations of quartzite, chert and carbonate skarn (within meta-argillaceous lithounit) and laminae of quartz–mica–chlorite within carbonate lithounit (Figure 2). These rocks are not included in the previous[5] stratigraphic scheme. Another interesting feature of the Sargipali deposit is the unusual Pb-rich mineralization with Cu association, which took place during deposition of the predominantly shaly sediments[4] (now represented by quartz–mica–chlorite schist, Figure 2) under reducing seawater environment. Lack of any contemporaneous oxidized facies around the ore body and the host quartz–mica–chlorite rock, the presence of thin pyrrhotite-rich bands in the deep mine levels, and the presence of graphite laminae in the ore and the gangue (Figure 3) collectively contribute to the idea of a reducing seawater environment. The present contribution, based on the study of petrogenesis of three rocks like marble, quartzite and skarn (which occur in the field as minor and impersistent lithounit), however, shows an evidence of oxidized environment during the formation of Proterozoic rocks in the area of study. Further, a geochemical evidence for the reducing condition is also being reported. Obviously, these findings are novel, since such type of data concerning this Proterozoic period remain rare worldwide.

The proposed programme has greatly benefited from the author's earlier work[4], in which a sufficient number of rock samples from various lithounits of the Sargipali area were analysed with respect to rare-earth elements (REEs). Out of these, three rocks, viz. marble, skarn and quartzite (which happen to be unusual on account of the mineral association presented in the following petrographic description), display a well-defined Ce anomaly. Therefore, the same is being reported and discussed as preliminary appraisal.

Furthermore, it may be imperative to point out that, although the analysed samples may not appear to be sufficient from statistical point of view, the REE trends represented by these analyses are quite logical and tenable in the light of the available literature. Hence, their utility is warranted in view of the fact[6] that sometimes, with the help of trace and rare-earth elements, geochemical modelling is possible even with one sample. With this assumption, a discussion on the above geochemical characteristics is made. But before this it is necessary to evaluate if the REE patterns of these

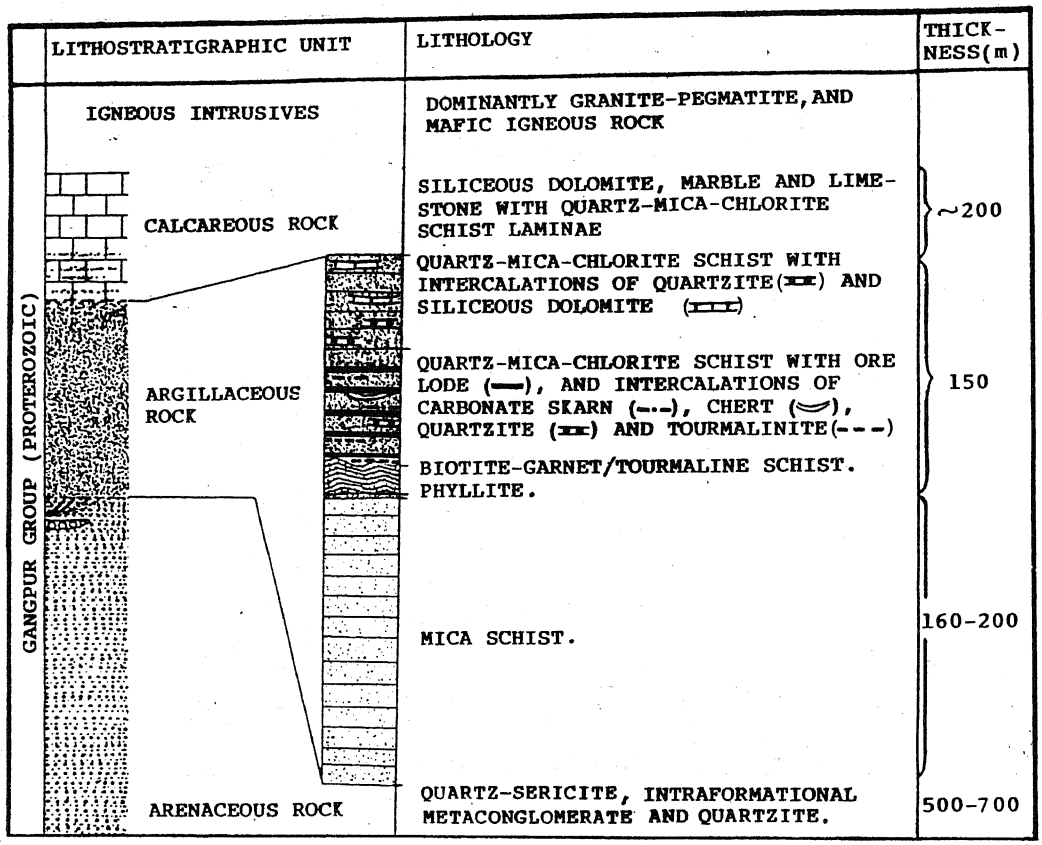


Figure 2. Generalized lithostratigraphy of Sargipali mine area (modified after Kar Ray[5]).

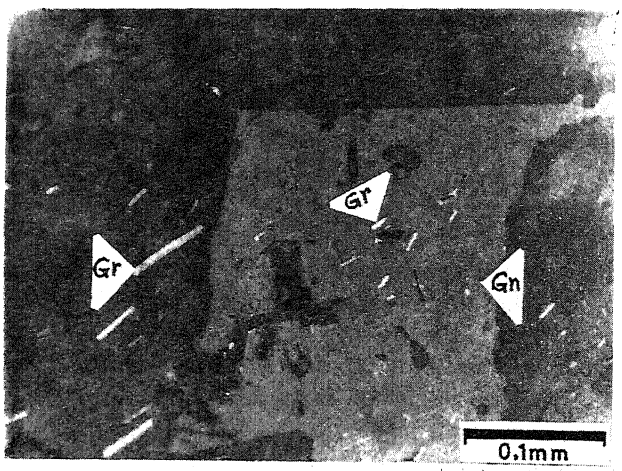


Figure 3. Occurrence of graphite in galena and gangue. Crossed polars. Specimen in oil. Gn = Galena, Gr = Graphite.

rock samples are at all representative of the original conditions of sedimentation, because the rocks have experienced multiple deformation, and also recrystallization over *PT* conditions of at least 600°C at 6 kb (ref. 4). To start with (i.e. the originality of REE pattern), the petrographic study is taken up. The *quartzite* sample, which has been collected from 'Mahikani' Hill, 4.6 km SSW to the prospecting mine (Figure 1), contains quartz and diopside as the major constituents; tremolite and zircon are minor while epidote is in the trace amount. Diopside occupies the intragranular space of quartz (Figure 4), containing noticeable amounts of zircon inclusions. Most likely, diopside in this rock represents a metamorphic equivalent of the initial impure carbonate which was mixed with initial quartz grains. It may, therefore, define a pure quartz sandstone, since in these rocks impure carbonate can occur as an irregular, scattered mass[7]. A provenance of this rock, as indicated by the total absence of feldspar and micas, could then lie chiefly with the quartzose sediment and/or the low-grade metamorphics of possibly quartzite. Petrographic study of *skarn* rock, which occurs as a thin laminate within quartz–mica–chlorite schist at 180 MRL of the underground mine reveals a calc-silicate mineral assemblage defined by diopside–tremolite–hornblende–chlorite–actinolite–calcite ± sphene. Here diopside is most abundant and, therefore, it can be a derivative of impure carbonate metamorphosed to amphibolite facies grade[8]. The petrographic study of *marble* (which is collected from its 2–3 inches thick concordant band in association with phyllite of the meta-argillaceous lithounit of the under-

ground mine at 120 MRL) reveals the groundmass of calcite to contain sometimes a minor patch of fibrous brucite. As a rule, this points to an original pure dolomite that upon metamorphism got decomposed and recrystallized to calcite and periclase with escape of CO_2 (ref. 9). In the presence of H_2O, however, periclase is transformed to brucite ($MgO \cdot H_2O$). This assessment, therefore, clearly agrees with the high MgO content of the analysed sample (Table 1). If it is so then the observed role of metamorphic decarbonation reaction is likely to disturb the reliability of the more sophisticated analytical data like $\delta^{18}O$ and $\delta^{13}C$ isotopes and REEs. Particularly, the isotopes of both $\delta^{18}O$ and $\delta^{13}C$ show a low value of -18.8 and $-11.01‰$, respectively (mean of three samples with respect to PDB), because CO_2 liberated could have been enriched in $\delta^{18}O$ as well as $\delta^{13}C$ relative to the carbonate[10]. On the other hand, compatibility of the REE data, particularly in relation to the precursor rock, is also a subject of examination, as follows in the ensuing discussion.

REE concentrations in the above rocks, obtained by inductively coupled plasma mass spectrometry with a precision of better than 7% relative standard deviation, are plotted as a REE pattern (Figure 5). Quartzite in this figure is rich in REE. This behaviour when considered in conjunction with the steep fractionated REE pattern may well represent a detrital origin of the rock. If this is true then, on the basis of the predominance of quartz, the rock can be conveniently assigned to sandy precursor; it is plausibly sandstone, because these are enriched in all REEs relative to chondrites[11,12]. Other samples, viz. marble and skarn, however, display depleted total REE with enriched heavy REE pattern (Figure 5). This behaviour not only corresponds to that of sediment but also suggests[11] that the rock is formed by the chemical precipitation process in marine environment. The reason behind the idea of marine hydrogenous origin lies in the fact that the above behaviour of REE is reminiscence of a low total concentration of REE with slight enrichment of heavy REE in seawater. In addition to this, it also appears that, although the rocks have evolved due to contact metamorphism, the magmatic water from the granitic stock did not introduce any large amount of REEs into the original rock compositions[13]. So, all these assessments may invariably support a general notion: the behaviour of REE during diagenesis and/or metamorphism indicate that there is little change in either the absolute abundances or the relative patterns.

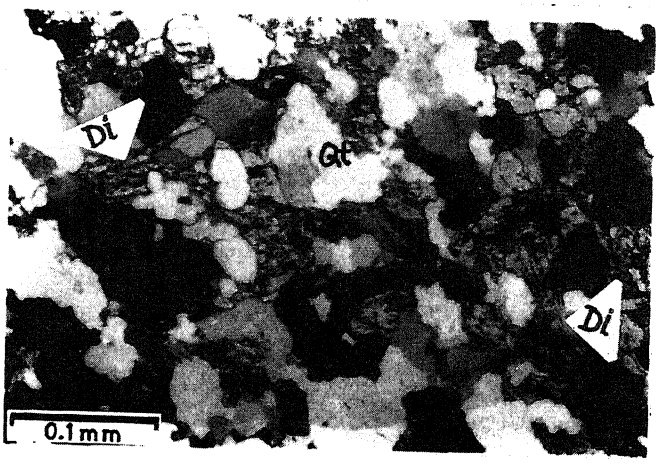


Figure 4. Photomicrograph of a rare quartzite showing diopside in the intragranular space of quartz grains. Crossed polars. Qt = quartz, Di = diopside.

Table 1. Major oxide composition in quartzite, skarn, marble and phyllite in (wt%)

	Quartzite	Skarn	Marble	Phyllite
SiO_2	82.74	48.73	1.02	60.12
TiO_2	0.03	0.07	0.00	0.53
Al_2O_3	0.54	6.02	0.04	18.40
Fe_2O_3 (T)	3.42	4.39	0.00	8.19
CaO	7.25	24.80	67.25	0.91
MgO	3.23	11.93	8.98	3.08
Na_2O	1.02	2.21	0.08	0.60
K_2O	0.39	0.35	0.01	3.62
MnO	0.12	0.42	0.37	0.23
P_2O_5	0.56	0.64	0.03	0.05
LOI	0.82	0.13	20.72	4.72
Total	100.10	99.69	98.50	100.45

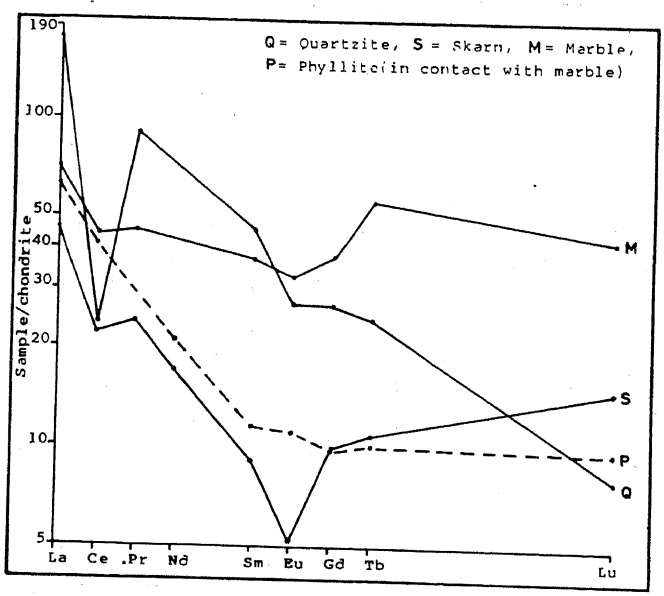


Figure 5. REE distribution pattern showing negative Ce anomaly in Sargipali quartzite, skarn and marble.

Nevertheless, in fulfilment of the reliability of the data, it is needed to bring out the factor governing minor Eu depletion in the studied samples (Figure 5). Critical to this, an effective way for removing Eu could have been the precipitation of $Eu^{2+}SO_4$. Such a mechanism is capable of removing most of the Eu while leaving most of the other REE[14]. If it is so, then at least most of the other samples formed in reducing conditions (e.g. quartz–mica–chlorite schist shown in figure 13 of Vishwakarma[4]) would have also yielded Eu depletion by this method. On the contrary, all these samples show positive Eu anomaly along with the steep fractionated REE pattern, and the features are actually, 'supplementary' evidence to charnockite provenance[4]. Here it may be annotated that by virtue of an evidence of the least Eu enrichment in the phyllite sample (Figure 5), which contains disseminated graphite, it cannot be invoked straightaway that the reducing conditions are prone to the formation of detectable Eu anomaly. Over and above, it is also not possible to propose removal of Eu anomaly due to remelting[15] or due to the influence of metamorphic fluid–rock interaction[16]. Likewise, the introduction of an oxidizing environment too has not a discernible effect on the behaviour of Eu[17]. Therefore, it seems that Eu anomalies in the samples of this study actually reflect an inherent property of their precursor. With this there remains no doubt that the REE patterns of the rocks are fairly representative of original sedimentation, and hence the observed feature of negative Ce–anomalies may be useful in providing an insight into the marine water condition during the formation of the rock[12, 17–20]. If this water is in an oxidizing condition, a smaller fraction of Ce will be trivalent, therefore resulting in Ce-depleted water condition and in turn Ce-depleted REE pattern in the marine sediments[19]. This scenario could have set in by the oxidation of Ce^{3+} under conditions of strong chemical complexation. Particularly, carbonate complexes of Ce^{4+} are the ones that control Ce in marine water (Möller, 1992, personal communication). Alternatively, according to experimental studies[21, 22], it is the oxidation of Ce^{3+}, into CeO_2 that causes Ce deficiency. This, therefore, corroborates the original idea of Goldberg[18]. The present study also reckons this observation in the light of the following. Phyllite (Figure 5) and quartz–mica–chlorite schist samples (Figure 13 of the earlier study[4]) both indicate a reducing environment of formation. So, for the obvious lack of oxygen, Ce removal by formation of CeO_2 could not take place, and as a result there remains normal La/Ce ratio, that is to say, normal REE pattern (Figure 5). Furthermore, the intensity of Ce anomaly observed in the present study is variable. Most likely, it is attributable to the quantum of the small fraction of Ce^{3+} which is left in an overall oxidizing marine condition.

1. Das Sharma, S., Srinivasan, R., Ahmad, S. M. and Patil, D. J., *Curr. Sci.*, 1994, **66**, 857–860.
2. Vishwakarama, R. K. and Ulabhaje, A. V., *Mineral. Deposita*, 1991, **26**, 26–29.
3. Vishwakarma, R. K., *Precamb. Res.*, 1995, in press.
4. Vishwakarma, R. K., unpublished Ph D thesis, Banaras Hindu University, Varanasi, 1990.
5. Kar Ray, M. K., *Indian Mineral.*, 1972, **26**, 37–47.
6. Knoper, M. W. and Condie, K. C., *Chem. Geol.*, 1988, **67**, 209–225.
7. Carozzi, A. V., *Microscopic Sedimentary Petrography*, Krieger, 1972, p. 9.
8. Nockolds, S. R., Knox, R. W. O. B. and Chinner, G. A., *Petrology for Students*, Cambridge University Press, Cambridge, 1979, pp. 399.
9. Bowen, N. L., *J. Geol.*, 1940, **48**, 225–274.
10. Shieh, Y. N. and Taylor, Jr. H. P., *J. Petrol.*, 1969, **10**, 307–331.
11. Graf, J. L. Jr., *Econ. Geol.*, 1977, **72**, 527–548.
12. Fleet, A. J., *Rare Earth Element Geochemistry* (ed. Henderson, P.), Elsevier, Amsterdam, 1984, pp. 343–373.
13. Jarvis, J. C., Wildeman, T. R. and Banks, N. G., *Chem. Geol.*, 1975, **16**, 27–37.
14. Kazimirz, J. B. and Madjar, J. B., U.S. Patent No. 3,092,449 (C 123–22), 1958, p. 3.
15. Taylor, S. R. and McLennan, S. M., *Precambrian Plate Tectonics* (ed. Kröner, A.), Elsevier, Amsterdam, 1981, pp. 527–548.
16. Danielson, A., Möller, P. and Dulski, P., *Chem. Geol.*, 1992, **97**, 89–100.
17. Taylor, S. R. and McLennan, S. M., *Geol. Soc. Am. Mem.*, 1983, **161**, 119–131.
18. Goldberg, E. D., *Oceanography*, American Association for the Advancement of Science, Washington, D.C., 1961, pp. 583–597.
19. Hu, X., Wang, Y. L. and Schmitt, R. A., *Geochim. Cosmochim. Acta*, 1988, **52**, 201–207.
20. Brookings, D. G., *Geochemistry and Mineralogy of Rare Earth Elements* (eds Lipin, B. R. and McKay, G. A.), Mineralogical Society of America, 1989, pp. 218.
21. Carpenter, J. H. and Grant, V. E., *J. Mar. Res.*, 1967, **25**, 228—238.
22. Zhenhua, Z., Nansheng, C., Zhensheng, D. and Zhnixin, Q., *Geochemistry*, 1984, **3**, 220–234.

ACKNOWLEDGEMENTS. Lt. Col. (Dr) V. Vishwakarma, Mrs Shail and Mrs Mohini encouraged me, and M/s Hindustan Zinc Ltd provided field work facility. Dr A. V. Ulabhaje, Prof. P. Möller and an anonymous referee reviewed the earlier write-up; Profs. R. K. Lal and S. G. Karkare helped in the petrographic study and major oxide determination, respectively; Drs S. M. Naqvi and V. Balram from NGRI and Mrs Kamlesh Agrawal from ONGC were courteous in helping me out with REE and isotope analyses, respectively; Dr N. Krishna Rao provided the facility for literature survey and ore microscopy; and a CSIR fellowship assisted me financially. I am grateful to all.

Received 25 July 1994; revised accepted 20 May 1995

Characterization of humic acids isolated from soils under mulberry vegetation

R. Kar, R. N. Dutta and S. K. Majumder

Section of Soil Science and Chemistry, Central Sericultural Research and Training Institute, Berhampore 742 101, India

Two different samples of humic acids were isolated from soils under mulberry vegetation. The isolated acids have been characterized by electrometric (potentiometry) titrations and visible spectroscopy with a view to probing their physicochemical properties. Attempts have also been made to correlate the results obtained from these experiments with coiling–decoiling behaviour as well as aliphatic–aromatic balance of the acids.

HUMIC acids are nothing but the natural polymeric acids having complex molecular nature[1]. Characterization of the same is necessary, not only for having some information about their structural arrangements, but also for the effective use of these substances in mulberry cultivation[2]. The principal features of humic acids – the polyelectrolytic behaviour and coiling–decoiling phenomenon – are probed in detail in the present paper by looking at different properties of two humic acid samples isolated from soils under mulberry vegetation located at two different agroclimatic conditions.

For isolation of humic acids, soil samples were treated with sodium carbonate, followed by acidification of the alkali-extracted material with dilute HCl to pH 2 (ref. 3). The purified humic acids were denoted as HA_b (humic acid sample isolated from the soil under mulberry vegetation of Central Sericultural Research and Training Institute, Berhampore, Gangetic alluvium of West Bengal, 24°4′ N–88°9′ E) and HA_k (humic acid sample isolated from the soil under mulberry vegetation of hill nursery, Kalimpong, West Bengal, 27°2′ N–88°17′ E).

Potentiometric titrations were carried out by using standard NaOH solution. Carboxylic and phenolic –OH group contents were reckoned from the acidity equivalents of the initial and final inflexion points of the titration curves[4]. The amount of fresh alkali required to attain a stable higher pH, following a fall in the final pH (on standing after initial titration), as well as the time elapsed for the same were also recorded for each sample.

The optical densities at 465 nm and 665 nm (ref. 1) for the humic acid solution (Na form) were noted with the help of a systronics spectrophotometer at two different pH values 7.0 and 9.0.

Table 1 records the carboxylic and phenolic –OH group contents and a pH difference namely, $\Delta pH = pH_{3/4} - pH_{1/4}$, where $pH_{1/4}$ and $pH_{3/4}$ refer to the values corresponding to one-fourth and three-fourths of the final inflexion points. The ΔpH values of the humic acid samples are greater than 0.954, which signifies the polyprotic nature of both the acid samples[5]. Again, the approximate charge density (number of charges per unit mass) of humic acid molecule is computed from the cumulative contribution of carboxylic and phenolic –OH contents. This, in turn, governs the 'adsorptive capacity' of humic acid molecules, which enables them to combine with metal ions, clays, pesticides, fertilizers and different organic compounds[6,7]. HA_k is found to be superior to HA_b in this aspect.

The amount of fresh alkali needed to attain a stable higher pH as well as the time required for the same have been reported to be correlated with the coiled nature of humic acid molecule[8–10]. Addition of fresh alkali for stabilizing the higher pH (when HA molecules were completely unfolded) indicates an amount of remainder hydrogen ions which are not accessible to alkali initially. This, in turn, is governed by the steric hindrance associated with a particular conformation of the molecules due to their coiled shape. Thus, for higher degree of coiling, greater amount of fresh alkali is

Table 1. Potentiometric information about HA samples

HA sample	Amount of COOH (meq · g^{-1})	Amount of phenolic –OH (meq · g^{-1})	Approximate charge density (Number of charges g^{-1}) × 10^{-20}	pH at first inflexion point	pH at final inflexion point	ΔpH	Conclusion
HA_b	1.63	4.14	34.73	8.72	9.48	4.74	Polyprotic
HA_k	4.03	6.05	60.68	7.96	8.66	4.33	Polyprotic

Table 2. Potentiometric stability and spectrophotometric analysis of HA samples

HA sample	Amount of fresh alkali required to attain stable pH (meq · g^{-1})	Time required to attain stable pH (h)	E_4/E_6 at pH		$\Delta E_4/E_6$
			7.0	9.0	
HA_b	0.87	23.00	3.32	2.75	0.57
HA_k	13.12	69.50	4.88	4.07	0.81

required. Again, the time required to attain a stable higher pH depends on the ease of decoiling, which is inversely related to the intramolecular hydrogen bonding. Formation of such bonds is often interrupted by the alkyl groups available in the humic acid molecule; this, in turn, governs the 'flexibility' of the molecule. Thus, poorer flexibility exhibits longer time requirement. From Table 2 it may be concluded that HA_k is more coiled but less flexible than HA_b.

The spectrophotometric analysis data have been recorded in Table 2. The E_4/E_6 values, being greater than unity at both pH values, establish that the absorbance by the aliphatic part (i.e. absorbance at 465 nm) is greater than that by the aromatic part (i.e. absorbance at 665 nm) due to a particular conformational arrangement of the aliphatic and the aromatic moiety[1,9].

Table 2 also shows the data on $\Delta(E_4/E_6)$ ratio, namely, $\Delta(E_4/E_6) = (E_4/E_6)_{pH-7} - (E_4/E_6)_{pH-9}$. The latter may be taken as an 'index of coiling' of the humic acid molecules owing to the fact that addition of alkali (NaOH) rendered the acid molecules 'unfolded under thermodynamic compulsion'[11]. This leads to the destruction of the tertiary structures (i.e. three-dimensional configuration) of the molecules following the release of the hidden aromatic parts, which thereby leads to a lowering of E_4/E_6 value, giving a positive value of $\Delta(E_4/E_6)$.The higher $\Delta(E_4/E_6)$ value of HA_k signifies that the latter is more coiled than HA_b. This has got close resemblance with potentiometric study discussed earlier. It may further be concluded that HA_k is more aromatized than HA_b.

1. Sarmah, A. C. and Bordoloi, P. K., *J. Indian Soc. Soil Sci.*, 1993, **41**, 642–648.
2. Inamatsu, K., *Bull. Seric. Expt. Stn.*, 1977, **27**(1), 1–96.
3. Adhikari, M. and Hazra, G. C., *Indian J. Appl. Chem.*, 1970, **33**, 322–328.
4. Schnitzer, M. and Khan, S. U., *Humic Substances in the Environment*, Marcel Dekker, New York, 1972, pp. 2–3.
5. Sturrock, P. E., *J. Chem. Educt.*, 1968, **45**, 258–259.
6. Kononova, M. M., *Soil Organic Matter*, Pergamon, Oxford, 1966, pp. 190–199.
7. Stevenson, F. J., *Humus Chemistry*, Wiley, New York, 1982, pp. 349.
8. Davis, H. and Mott, C. J. B., *J. Soil Sci.*, 1981, **32**, 393–397.
9. Saha, P. B. and Sanyal, S. K., *J. Indian Soc. Soil Sci.*, 1988, **36**, 35–42.
10. Kar, R. and Sanyal, S. K., *J. Indian Chem. Soc.*, 1988, **65**, 834–837.
11. Sanyal, S. K., *Chem. Edn.*, 1984, **1**, 14–18.

Received 28 December 1994; revised accepted 22 May 1995

Research Accounts

Current Science intends to start a new feature entitled 'Research Accounts'. Articles in this category are intended to be personalised reviews of research from the authors own laboratory, based on a body of published work. The articles must provide appropriate background to the area in a concise introduction, which should also serve to place the author's work in proper perspective. A concluding section may address future trends and directions in the area. Unpublished results may be included only when absolutely essential and should be kept to a bare minimum. Articles in this category may be about 20 double spaced typed pages; A4 size (8–10 printed pages) and should include appropriate illustrations. References to the literature should not number more than about 60. Articles should address a broad readership. While most articles in this category will be solicited by the editors, unsolicited contributions will also be considered. All articles will be reviewed by two independent referees. Authors are encouraged to suggest appropriate reviewers.

— Editors

Contributions to Operator Theory and its Applications. T. Furuta, I. Gohberg and T. Nakazi (eds). Volume 62 in the series Operator Theory: Advances and Applications, Birkhäuser, P.O. Box 133, elt 4010. Basel. 1993.

The name Tsuyoshi Ando is well known to research workers in functional analysis, operator theory, linear algebra and the mathematical theory of networks and systems. This volume consists of thirteen papers in these areas dedicated to Ando on his sixtieth birthday. A brief biography of Ando and his list of publications is also provided.

Since the papers in this book are on diverse problems, the best way to review it might be to give the reader a general idea of some of the topics covered, with particular reference to the work of Ando.

Let $\mathcal{H}$ be a Hilbert space contained in a larger Hilbert space $\mathcal{K}$. An operator T on $\mathcal{K}$ is called a *dilation* of an operator A on $\mathcal{H}$ if T can be written as

$$T = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

when $\mathcal{K}$ is decomposed as a direct sum of $\mathcal{H}$ and its orthogonal complement. The operator A is then called a *compression of T*. A *dilation theorem* is a statement that certain operators can be dilated to 'nicer' operators on a bigger Hilbert space. For example, if A is a contraction ($\|A\| \leq 1$) on $\mathcal{H}$ then

$$U = \begin{bmatrix} A & (1-AA^*)^{1/2} \\ (1-A^*A)^{1/2} & -A^* \end{bmatrix}$$

is a unitary operator. This is a generalization of the elementary fact that if a is a real number and $|a| \leq 1$ then

$$\begin{bmatrix} a & \sqrt{1-a^2} \\ \sqrt{1-a^2} & -a \end{bmatrix}$$

is a rotation matrix. Thus, *every contraction has a unitary dilation.*

If A is a contraction then so are its powers A^n, $n = 1, 2, \ldots$. Each of them will have a unitary dilation. It is somewhat surprising that there exists a single unitary operator U such that U^n is a dilation of A^n for all n. This is a famous theorem proved by, Sz.-Nagy in 1953. (The discrete variable n here can be replaced by a continuous variable t: every contraction semigroup can be dilated to a unitary semigroup).

Results such as this are typical in dilation theory, now a major subject. One of the famous results of Ando, proved in 1963, says that a *commuting pair of contractions can be dilated to a commuting pair of unitaries,* i. e., if A_1, A_2 are commuting contractions, then there exists commuting unitaries U_1 U_2 such that for all positive integers n, m, $U_1^n U_2^m$ is a dilation of $A_1^n A_2^m$. In 1970 it was proved by Parrott that an analogue of this for three or more commuting contractions is false. The *commutant lifting theorem* to which this line of work ultimately led has been found to be very useful in the theory of linear systems.

Let P_1, P_2 be two projection operators in a Hilbert space. Long ago, von Neumann observed that the infinite product $(P_1 P_2)^n$ converges to the projection operator whose range is the intersection of the range of P_1 and P_2. In 1961, Burkholder and Chow proved that the same is true for any finite number of projections $P_1, P_2, \ldots, P_k$, and when the Hilbert space is L^2 and the projections are conditional expectations then the convergence takes place not only in the strong (L^2) topology but also pointwise, almost everywhere. In a different direction, Amemiya and Ando proved in 1965 that if $T_1, \ldots, T_k$ are contractions then any *random* product $T_{i_1} T_{i_2} T_{i_3} \cdots$ of elements chosen from $\{T_1, \ldots, T_k\}$ converges weakly. Such results are closely related to problems in ergodic theory, dynamical systems and computer tomography.

The arithmetic, geometric and harmonic means of positive numbers are familiar objects. Lots of interesting difficulties arise in defining similar means for positive Hilbert space operators, because of the noncommuting nature of the variables. Such means are of interest in quantum theory and also in the theory of electrical networks. For example, the parallel sum is just one half of the harmonic mean. Ando has studied such objects in great detail. In a paper with Kubo in 1980, he gave an axiomatic characterization of such means. In a very interesting paper (which drew this reviewer to his work) Ando used these ideas to give a simple proof of a theorem of Lieb that established the Wigner–Yanase–Dyson conjecture about the concavity of entropy in quantum mechanics.

From the late seventies, much of Ando's work has been on matrices, specially on matrix inequalities. He has mastered the techniques of majorization and the theory of operator monotone functions, and used both of them in spectacular ways to prove delicate and elegant inequalities for operators and matrices.

The book under review has papers on many of the problems to which the work mentioned above leads. It would be of interest to anyone working in these areas.

RAJENDRA BHATIA

Indian Statistical Institute
New Delhi 110 016, India

Science in Society: Some Perspectives. Yash Pal, Ashok Jain and Subodh Mohanti eds. NISTADS, Dr K. S. Krishnan Road, New Delhi 110 012. 405 pp. Rs. 400.

This handy anthology starts with a magnificent introduction from Prof. Yash Pal; he leads off the discussion with a characteristically brilliant and thought-provoking statement, as arresting as it is pithy. The lead sentence 'the difference between science and society and science in society is very important for me' sets the mood and the direction for all the essays in the book although not all of them reach the standard of clarity of thinking or of focused purpose that illuminate Yash Pal's lead. The book demonstrates congruence of views and should provide the basis for concerted action, review and mid-course corrections in policy.

One difficulty that an average reader will find is that the book bears no explicit reference to the year in which these essays have been prepared and no indication if some of the initiatives mentioned are under implementation or if they are mere expressions of hope (the case of INFLIBNET is a clear case in point). Some of the articles would have been more convincing if they had included some idea of costs to the user. An epilogue to show how far the movement initiated by Yash Pal in the 77th Indian Science Congress has developed momentum would have been useful.

Science in society

Yash Pal leads the way with a complete statement of the case for 'science in society'. Three quotes would give the reader an insight into what to expect: 'science should not be treated as an additional activity or a tool ... instead it should be treated as part of the culture of society, integral to our thinking ... connected with our dreams, connected with concepts of ethics, beauty and spirituality', 'if we make science as a part of our culture they (benefits) would all come, much more naturally, more creatively and *more in tune with our living, not as an assault from outside*', 'in science and technology you cannot imbibe the true spirit of the craft *unless you do some things on your own*, do your own inventing'. Yash Pal covers a wide territory: the value of innovations to true development, the imperative need to invest in education if S & T (science and technology) R & D is to be supported, an equal emphasis on investments in S & T and in science education. There is a touch of humour in his development of the relationship of ethics and law in an age of 'technological exuberance': he says that 'it was not unethical 500 or 1000 years ago to throw your muck into the Ganges ... because Ganga had sufficient powers to purify it'.

Many of the contributors clearly link their contributions to 'science in society', the absorption of science as an integral part of our culture. The contributions can be divided into three categories: those with a philosophical flavour representing *cultural absorption*, those with plans and ideas to *involve the population* in developmental activities as a practical illustration of how 'science in society' could operate, and those which demonstrate *vigour in scientific thought and practice* and induce in the reader a sense of optimism that Yash Pal's vision can indeed become a reality.

Cultural absorption

Pushpa Bhargava and Chandana Chakrabarti see a commonality in the origins of the creative urge in scientific and artistic works, describe practitioners of art and science as seekers of beauty and assert that both (art and science) 'are becoming increasingly interdependent in terms of materials, ideas and ideals'. In a fine essay 'The symphony of symmetries', Babu Joseph points to developments in the study of open systems which 'include the whole living world as well as most nonliving things'. His development of ideas on 'spontaneous symmetry breaking' as the mechanism of growth and decay should be fed into the curriculum for science at several levels with increasing depth of presentation. Ashok Parthasarathi's coupling of nonalignment, self-reliance and innovation gives practical shape to 'science in society' (particularly Yash Pal's call for 'doing something on your own') and brings policy-making into focus. In particular, his declaration that 'while it is true that S & T intersect at many points, the intersections are less and less matters of accident and more and more matters of design in terms of policies and programmes' deserves the basis for nonpartisan policy-making. Another very interesting articulation of the concept of 'science in society' is the use of energy as a measure of programme content (C. V. Seshadri), which leads to sustainable development through the blended use of resources combining development with avoidance of environmental degradation. D. P. Chattopadhyaya's article draws attention to the two faces of technology: estrangement and enlargement, and emphasizes the need for an awareness of the consequences of the use of technology. His is a timely call for 'workers and thinkers to control technology and institutions more effectively than they can control us'. A similar warning on the impact of technology on society comes from V. Sudarsen, who calls for alternative paradigms of development citing the irony of continuity in the pursuit of an imperial forest policy even after independence. He warns against the disregard for social consequences and asserts that 'technological progress is not necessarily axiomatic for development processes'. Ashok Jain echoes this thought with 'environmental crisis is a sign of saturation of the conventional paradigm'.

Precept and practice

Jayant Narlikar's fine essay 'Education through astronomy' will find an echo in many whose introduction to astronomy was wholly through gymnastics in integral or differential calculus. Equally well-correlated to the theme of the book is V. G. Bhide's note on interuniversity centres; he is the lone author tracing lack of a clear career path (salaries included) as one major cause of student apathy. He also relates the government-inspired and organized techniques to the dual demands of the 'product of science' and the 'process of science'. His essay is the kind of core strategy that matches the vision of 'science in society', catching them young as it were. Vinod Raina's article describing a programme of the 'People's science movement' is an excellent example of how a concept can be translated into practice and harmonized with the circumstances in a given location. This article should give a fillip to the emergence of more initiatives of the kind. M. S. Swaminathan's contribution 'Biotechnology and ecological security' is scholarly and practical, introducing us to the concept of 'ecological security'. His article is an outstanding example of how a versatile mind can be ahead of formalized thought. His suggestions and ideas on directions of development give practical content to the concept of 'science in society' through the fusion of physical and social sciences to produce worthwhile and commercially viable goals.

Excitement of science

The articles on scientific projects (earthquakes, monoclonal antibodies) reveal the capabilities of vigorous research in India but the authors would have expanded their audience reach if they had explained the techniques and results in layman terms. The sense of excitement in science needs to reach young minds in as many directions as possible. For example, apart from publications intended for professional journals exclusively dedicated to research material, can we not think of articles on S & T that describe briefly, or at least enumerate, the different technologies on which the conduct of an item of research or production of specified goods and services depend? Such articles would show young enquiring minds what career path options in S & T may be available for them to choose from, and also infuse in them a sense of team-orientation in S & T work.

A strong public opinion is the proper basis for the percolation of the 'science in society' concept and its acceptance. The editors of the book have rendered a great service by bringing out this volume

in order to catalyse this process. Broad-based initiatives which bring the concept into practice using local resources would be necessary. This also suggests that research effort is needed for the evolution and adoption of organizational structures suitable for the task on hand in the region where the initiative takes shape. The apparent apathy in 'science in education' can be overcome only through multiple techniques and structures. Perhaps the disavowal of the editors about the book not being a 'cogent articulation of various tenets' clearly recognizes the need for further effort. The mention of a 'think tank' might draw cynical expressions of elitism but the time is opportune for one or more 'think tanks' to address these issues and develop a menu of options to carry Yash Pal's lead concepts onward; local groups could choose their strategy from this menu for use as suits circumstances in their area. The ideas sprouting from the studies will motivate policy-making at governmental and other levels. This should lead to the gradual infusion of science in national culture. Thoughtful leaders of public opinion (of the kind who have contributed to this book) should also find out how the stated objective could be reached in an environment of limits to the efficacy of governmental patronage.

R. PARTHASARATHY

S-23 Golden Enclave
Airport Road
Bangalore 560 017, India

Annual Review of Cell Biology 1994. J. A. Spudich, S. L. McKnight, R. Schekman, eds. Annual Reviews Inc., Palo Alto, California, USA. vol. 10. 478 pp. Price: USA $ 49; elsewhere $ 54.

Cellular biology continues to grow and diversify, encompassing the broad areas of biochemistry, genetics and molecular biology, in order to obtain an understanding of the inner working of the cell in molecular terms. These endeavours have now been extended to the study of whole life forms, as topics in developmental biology find a larger representation in this volume of *Annual Review of Cell Biology*, paving the way for a formal merger of these two disciplines from 1995 onwards, when the name of the series will change to *Annual Review of Cell and Developmental Biology*.

The pathways that control cell growth and differentiation, in mammals in particular, are undoubtedly complex. Recent progress in this field has been remarkable and five reviews are devoted to advances in the broad area of cell regulation. In a lucid article by D. J. Riley and coworkers on the retinoblastoma protein, the properties of this key regulator of cell division have been discussed. In the normal cell cycle, activated retinoblastoma protein restricts further cell proliferation. However, binding of oncogene products to this protein can block its activity and lead to unscheduled cell division, resulting in malignancy. U. Siebenlist and coworkers have reviewed the structure and functions of the transcription factor NFκB family of proteins, which serve to coordinate the rapid induction of a range of defence genes in the immune system. A unique feature of NFκB is its rapid translocation from the cytoplasm to the nucleus in response to extracellular signals, and the complexities of this process have been analysed in depth.

Signal transduction pathways in the cell basically depend on the fine tuning of phosphorylation of specific proteins by kinases and phosphatases. P. van der Geer and others have given a detailed account of a major class of transmembrane receptors called the receptor protein–tyrosine kinases. Studies on these kinases have contributed significantly to an understanding of how cells are regulated by extracellular signals. S. Shenolikar has written a comprehensive review on the protein serine/threonine phosphatases, highlighting their physiological functions and cellular regulation. A group of proteins that play a crucial role in signal transduction pathways are the GTPases. T. A. Glomset and C. C. Farnsworth have documented the different posttranslational modifications found in these GTPases (such as farnesylation of the carboxy terminus and acylation of the amino terminus) and presented evidence for their role in determining the interactions between the GTPases and the cell membranes. Additional studies are required before the functions of these modifications are firmly established.

In the area of intracellular sorting of proteins, which found exhaustive coverage in last year's volume, there is a single review on protein targeting to the endoplasmic reticulum by P. Walter and A. E. Johnson, which emphasizes the common steps in targeting in eukaryotes and prokaryotes. However, the mechanism of translocation of the nascent protein across the endoplasmic reticulum remains less well-defined.

The cytoskeletal framework of the cell, for long a favoured topic for cell morphologists, is emerging as a dynamic and complex structure, subject to a variety of regulatory phenomena. Five reviews are devoted to recent developments on the structure and function of cytoskeletal and associated proteins. A. Hall has contributed a highly topical review on the regulation of the actin cytoskeleton in response to extracellular signals through the ras family of small GTP-binding proteins. A comprehensive review on the three-dimensional structures of 12 actin-binding proteins has been provided by T. D. Pollard and others. E. L. F. Holzbaur and R. B. Vallee have discussed the recent findings on the structure and functions of dyneins, one of the largest and most complex motor proteins. A major class of cytoskeletal proteins, the microtubules, are involved in cellular elongation and cytokinesis. Their role in cellular patterning in the growing plant has been elegantly discussed by R. J. Cyr, who has also presented a hypothesis for the mechanism by which microtubules bring about cellular elongation.

Vertebrate limb development has been one of the best-studied systems for understanding cell and tissue patterning, mechanisms of innervation, cell differentiation, angiogenesis and programmed cell death. C. Tickle and G. Eichele have reviewed the embryological data on limb patterning, and presented recent work that has begun to unravel the molecular responses to signalling molecules involved in the process.

Intensive research on the mechanism of Alzheimer's disease has led to the identification of cerebral deposits of the amyloid β protein as the causative agent of the disease. D. J. Selkoe has reviewed the cell biology, in particular the intracellular processing, of the amyloid β protein precursor. Several mechanisms have been proposed for the eventual imbalance between amyloid β protein formation and clearance which results in

amyloid deposits. Pharmacological intervention in the disease process is being attempted at various points along the amyloidogenic pathway, and may have the potential of being used to slow down the gradual accumulation of amyloid deposits that accompanies normal aging of the human brain.

This volume of the *Annual Review of Cell Biology* truly 'unifies a diverse and rapidly expanding field'. The reviews are selective and critical in format rather than exhaustive. An essential book for all researchers in cell biology.

VEENA K. PARNAIK

Centre for Cellular & Molecular Biology
Hyderabad 500 007, India

Environment and Adaptive Biology of Plants. Professor David N. Sen Commemoration Volume. D. D. Chawan, ed. Scientific Publishers, 5A, New Pali Road, Jodhpur 342 001, India. 1995. Price: Not known. 322 pp.

Environment and adaptive biology of plants is one of the topics of past, present and future scientific interest. Much of the progress we have made in crop selection and breeding have stemmed from basic understanding of plant adaptation to environmental conditions. The success of introduced agricultural, horticultural and timber plants all over the world, ever since humans began transporting them, has also been a result of our knowledge of plant adaptability to different environment.

Current interests in plant adaptive biology worldwide are in the areas of (1) *in situ* protection and management of endangered plants, (2) sustaining productivity in crop plants, (3) identifying locally evolved resistance to pests, diseases and other environmental stresses, (4) reclamation and revegetation of degraded landscapes and (5) management of increasing salt and water stresses in coastal and semi-arid zones.

Considerable progress has been made in the above fields of environmental science both in India and abroad. Our knowledge of dioecious plants (especially in the genus *Garcinia*) being capable of reproducing apomictically on isolation is helpful in the management of threatened species. Case studies of such plants shed light on survival strategies that isolated plants adopt at the expense of genetic variability. Land races that are locally adapted are in much demand today for crop improvement and sustainable agriculture. Most recent interest has been in understanding the mechanisms plants have adopted in dealing with salt stress. Of special importance in this regard are the estuarine plants, commonly referred to as mangroves. Mangrove plants in the genus *Rhizophora* are known to localize the salt in the root zone itself. These plants, which are also capable of withstanding flooded conditions, can be the answer to ecological problems that might arise in the event of the anticipated sea level rise. Research on the physiology and genetics of mangrove plants is under way, especially in an effort to identify 'plus' trees (supertrees which can withstand considerable amounts of environmental stress). A major programme in this field of interest has been on at the M. S. Swaminathan Research Foundation, sponsored by the Department of Biotechnology, Government of India.

The book under review is the most unusual blend of biography, science, documentation and advertisement. The book is divided into six parts, viz., (1) Ecology and environment management, (2) Adaptive biology, (3) Biology of saline plants, (4) Weed biology, weed management and allelopathy, (5) Desertification and agroforestry and (6) Environmental pollution. Two to five papers have been included under each of these major parts. These papers vary in style from being introductory to highly structured experiments and results, making reading rather uncomfortable. Papers have not been appropriately grouped. For instance, Part 1 includes an ethnobotanical document which merely lists plants used by natives in the Narmada valley. Similarly, in Part 4, one paper provides a detailed list of weeds associated with crop fields of Madhya Pradesh without any reference to management.

The strength of this book, however, rests in Part 3, which discusses a topic of current scientific interest. It appropriately includes papers on mangrove plants and those in salt-prone deserts of India. Although the discussions fall short of being comprehensive, the papers are more in tune with the broad title of the book than the rest.

The book, in general, fails to address current issues relating to the environment and adaptive biology of plants. All detailed discussions included have emerged from laboratory studies paying little attention to field conditions. Most papers are not even state-of-the-art treatises of the concerned science, not even of that in India. To a serious student of plant adaptive biology, the book can well prove disappointing.

R. J. RANJIT DANIELS

M. S. Swaminathan Research Foundation
3rd Cross Street
Taramani Institutional Area
Madras 600 113, India

A pioneer in modern plasma physics

An obituary of Hannes Alfvén

Professor Hannes Olof Gösta Alfvén, a renowned space physicist and a prime contributor to the field of magneto-hydrodynamics, died on 2 April 1995 at his home in Djursholm, Sweden, at the age of 86. With his passing away, the world of physics has lost a highly creative and original physicist. He can justifiably be called one of the pioneers of modern plasma physics – a field of study that finds applications in a wide range of areas such as cosmology, astrophysics, controlled thermonuclear fusion, radio communication, etc. Alfvén's contributions and insights into the field, especially during its infancy, are many and bear the stamp of his remarkable physical intuition. He developed very powerful tools which greatly aided the analysis of complex physical phenomena and advanced the field in giant steps.

An early example of such a powerful idea is that of the concept of frozen-in flux, where the plasma is considered to be tied on to the magnetic lines of field that pass through it. Alfvén discovered this theorem in the early 1930s and used this simplifying tool to great advantage in solving many space physics problems. In fact, so successful was this paradigm that it began to be used indiscriminately by many later workers in the field, who applied it to situations where its validity conditions were often violated. Alfvén expended considerable effort in educating the scientific community about the proper validity conditions of this theorem. Another simplifying notion that Alfvén invented was the guiding centre approximation, by which the complex orbit of a charged particle in a magnetic field could be calculated in a very simple manner in terms of the motion of the centre of the Larmor motion and its drifts. The physical notion underlying this idea is that of the equivalent magnetic moment associated with the spiral motion of the charged particle. This concept helped him understand many complex phenomena associated with auroras, magnetic storms, cosmic ray generation in the solar system and in the interstellar space. In the laboratory, this concept helped in the development of some early practical devices like the trochotron, a high vacuum switching tube, and in later years the idea of the invariance of the magnetic moment was at the heart of the plasma confinement in mirror machines. Today some of these tools invented by Alfvén have become such standard repertoire of plasma physicists that we hardly pause to contemplate their power and simplicity or to imagine the difficulties of working without their aid.

His most well-known discovery – that of hydromagnetic waves, which are commonly called Alfvén waves – is another prime illustration of the power of his physical intuition. In thinking about the problem of sunspots, he reasoned that they consisted of magnetic fields embedded in the highly conducting solar plasma. These magnetic fields must arise from currents flowing in the plasma and the mutual interaction between them must make both the currents and the fields to move. Thus on purely physical grounds he worked out the theory of this movement – that of an electromagnetic wave – propagating through a highly conducting medium. At the time when he published his theory in *Nature* in 1942, nobody believed it. It took many years for scientists to accept this idea. The common notion at the time, from the well-known Maxwell's theory, was that an electromagnetic wave could scarcely penetrate a conductor. Thus, Alfvén's proposition of an electromagnetic wave propagating in a perfectly conducting medium with no attenuation seemed physically impossible. It was only in 1948 when Alfvén visited the United States for the first time and lectured about his work at a number of places that the correctness and significance of his work was recognized. According to an apocryphal story, when Fermy heard his lecture at the University of Chicago, he nodded his head and said, 'Of course'. The next day the entire world of physics said, 'Oh, of course'.

Alfvén was also the first person to suggest the existence of large-scale magnetic fields in the galaxy. This was soon after it was discovered that cosmic rays were charged particles and not gamma rays filling the entire universe. This spurred Alfvén to create a self-consistent picture of electrical currents flowing in a low-density plasma pervading the galaxy, giving rise to large-scale magnetic fields and the cosmic ray particles moving in spiral orbits about these fields. These ideas, which are well proven and accepted today, were way ahead of their times in the late 1930s when interstellar space was supposed to be a vacuum and could not support any currents. In 1939, he also proposed a remarkable theory of magnetic storms and the aurora in which he introduced several novel concepts, including that of the guiding centre approximation, the concept of electric field drift for plasma within the geomagnetic field and the concept of a ring current formed from quasi-trapped radiation. He preceded Fermi in suggesting a mechanism for cosmic ray acceleration and also provided the right tools for understanding phenomena such as the Van Allen radiation belts.

Alfvén's intuitive approach was often in conflict with the prevailing style of doing science with idealized mathematical models which could be solved exactly. Thus, he often faced difficulty in publishing his papers. But his first book, *Cosmical Electrodynamics*, published in 1950, in which he collected much of his early work, had a great impact in the development of space physics and plasma physics. This development also vindicated many of the fundamental concepts introduced by him. Thanks to his pioneering contributions, the theory of magnetohydrodynamics has today become an established and well-developed branch of physics and is widely used in the field of astrophysics and fusion physics.

Alfvén was acutely aware of the importance of technological developments leading to the creation of new fields for scientific investigations and believed that 'the centre of gravity of the physical sciences is always moving'. He was thus an active supporter of newer space experiments and advances in fusion devices. He also believed in small-scale laboratory experiments to test basic concepts and promoted actively the development of such programmes in the institutions that he was associated with. Some of them carried out remarkable fundamental

experiments in plasma physics, often in the form of verification of some of his own ideas. For instance, in developing a comprehensive theory for the evolution of the solar system, Alfvén intuitively proposed the concept of a critical relative velocity between a neutral gas and a magnetized plasma. Beyond this velocity the gas would ionize rapidly and the resulting plasma cloud would form the starting point for the process of condensation and accretion, leading to the formation of secondary bodies like planets and satellites. The value for the critical velocity was proposed to be $(2eV_i/M)^{1/2}$, where V_i and M are the ionization potential and mass of neutral atom. This limit has been experimentally verified in a number of experiments, including some carried out at the Royal Institute of Technology in Stockholm under his active encouragement. Likewise, he inspired a series of experiments and developments on the Marshall gun – a plasma device to create a minimum magnetic energy plasma configuration which he had theoretically proposed. This concept in the advanced form of a Spheromak configuration is today considered to be an attractive fusion confinement geometry. His original idea of intense electric field structures in space led to a large body of theoretical and experimental investigation into the so-called potential double-layer formations. Alfvén also helped inspire some early plasma physics work in India when he visited the Tata Institute of Fundamental Research, Bombay, in 1958 at the invitation of Homi Bhabha and gave a series of lectures on plasma physics to the then fledging plasma physics group.

Not all his ideas are accepted even today. He has put forward his own theory of cosmology, opposing the big bang model, and leaning towards the cosmology of Oscar Klien, with a universe oscillating between expansion and contraction. He believed that the universe was formed as an ambiplasma (i.e. with equal amounts of matter and antimatter) and suggested clever mechanisms for their segregation into separate worlds. He also emphasized the importance of dust in the universe and talked long time ago about 'dusty' plasmas. Cast in the role of an iconoclast, during his early struggles with the established scientific views in space physics, Alfvén continued to question the established doctrine and always fought for his own 'wild' ideas. In the words of his friend and associate, Gustaf Arrhenius, Alfvén was a 'gentle wild man'. He was also an active campaigner against nuclear warfare and participated in the East–West 'Pugwash' conferences that sought to prevent such confrontations. Originally a strong promoter of nuclear power generation, he later grew apprehensive about the long-term safety of reactors and sought to highlight the problem of nuclear wastes. He continued to be active in research till fairly recently, sharing his time between the Royal Institute of Technology and the University of California in San Diego.

Alfvén was the recipient of several awards, including the Nobel Award in physics in 1970 (shared with Louis Neel), the Lomonosov Gold Medal of the Soviet Academy of Sciences, the Franklin Gold Medal of the Franklin Institute in Philadelphia and the Gold Medal of the Royal Astronomical Society, UK. He was a member of the National Academy of Sciences in the United States and a Fellow of the Royal Society in Britain. His books included: *Cosmical Electrodynamics*, 1950; *On the Origin of the Solar System*, 1954; *Cosmical Electrodynamics: Fundamental Principles* (with C. G. Fälthammar), 1963; *Worlds–Antiworlds*, 1966; *Structure and Evolutionary History of the Solar System* (with G. Arrhenius), 1975; *Cosmic Plasma* (monograph), 1981. He also wrote several popular scientific books, many of them in collaboration with his wife Kerstin Maria Erikson and often assuming the pen name of Olof Johannesson.

Alfvén has given the world of physics a rich legacy of ideas and tools. They have helped us reach out and understand many complex phenomena in the sun, the stars and the space beyond. They are also the basis on which one hopes to create the first star in the laboratory – and harness the energy of fusion reactions. Such is the power and reach of the field he helped create and which reflects in a striking way the true measure of his amazing genius.

ABHIJIT SEN
P. K. KAW

Institute for Plasma Research
Bhat
Gandhinagar 382 424, India

Remembering Hannes Alfvén and his work

'You know I love your country and I do Yoga everyday', Professor Hannes Alfvén was telling me this in a conversation in Gothenburg, Sweden, at a Plasma Physics conference in 1988. He further added, on seeing the book *The Alfvén Wave*, the first book on this topic written by Akira Hasegawa from Bell Laboratories, NJ, USA and myself and dedicated to him, 'You both seem to know more these waves than myself'. This was the highest complement coming from Hannes Alfvén, who discovered the Alfvén waves in 1942.

He can be justifiably considered as the father of space plasma physics and astrophysical plasmas. He originated many key theoretical ideas which opened a new vision in understanding of the plasma universe. Hannes Alfvén received his Doctorate from the University of Upsala in 1934 and was appointed as Professor of electromagnetic theory and electric measurements at the Royal Institute of Technology, Stockholm, in 1940, at the age of 32. From this date to almost the end of his life, Alfvén was an active scientist of this laboratory. His vigorous scientific and administrative activities led to the creation of several new departments within the Royal Institute of Technology. Three of these, Plasma Physics, Fusion Plasma Physics and Accelerator Technology Departments now constituted a separate entity called the Alfvén Laboratory. His own chair evolved correspond-

ingly into electronics in 1945 and plasma physics in 1963.

From very early age, as a young boy Hannes Alfvén was intrigued by radio sets and radio electromagnetic waves. His vision about magnetic and electric fields and interaction with matter in the universe led him to discover the entirely new kind of waves in conducting fluids interacting with magnetic fields, which are now known as Alfvén waves. These waves have become a foundation stone in plasma physics and its applications to space physics, astrophysics and thermonuclear fusion plasma physics. In order to understand the theory of sunspots, Alfvén in a now famous letter to *Nature* gave a very simple mathematical formulation describing the mutual interaction between electromagnetic fields and conducting fluid motion and found that this interaction can give rise to a new wave phenomenon called electromagnetic–hydrodynamic waves.

Alfvén waves, as these waves are now known are low-frequency transverse waves propagating along the magnetic lines of force. Unlike sound waves, these are choosy about the direction of propagation. Closely related to the Alfvén waves is the concept of 'frozen in magnetic fields', also due to Alfvén. This concept simplifies greatly the physical reasoning about plasma phenomena but has its limitations. Using this concept it was very easy for Alfvén to draw an analogy with the theory of stretched strings. The magnetic tension arising due to parallel component of the Lorentz force along the magnetic field lines can lead to the possibility of transverse waves along the lines of force, with velocity ϑ_A given by

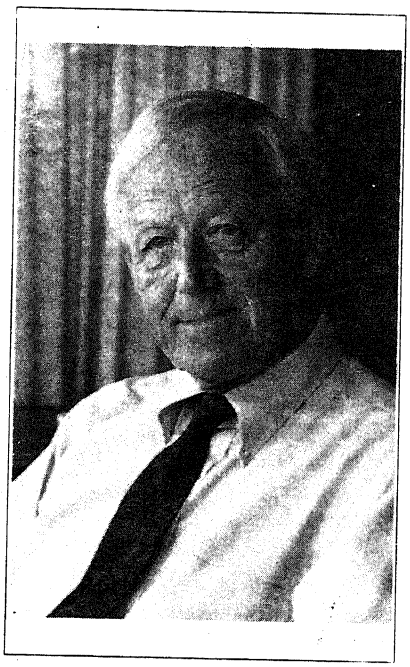
$$\vartheta_A = B_0/(\mu_0 \rho)^{1/2}$$

where ρ is the mass density of the fluid on which the field lines are imbedded.

The 'frozen in' concept, though simplifies greatly the physical reasoning about plasma phenomena, has its limitations. Alfvén himself cautioned the plasma physicists in using this concept without proper justification.

Incredibly, it took several years before the results arrived at by Alfvén were taken seriously. The experimental verification of Alfvén waves came still later due to technical difficulties. With the advent of space age now we have numerous number of satellite and spacecraft observations pointing out the omnipresence of Alfvén waves in space physical systems. The theoretical studies point out how various natural phenomena in nature can be explained as manifestation of the presence of Alfvén waves. In recent years the study of Alfvén waves in inhomogeneous magnetic fields have led to a considerable understanding e.g. of solar coronal heating, bending waves in galaxies, the characteristic features and origin of ULF waves in the magnetically organized systems, the magnetosphere, particle acceleration in auroras and so on.

The other natural phenomenon which attracted the attention of Hannes Alfvén



was the aurora. As we know auroras are a spectacular display of luminous radiation in the sky near polar regions. Auroras as we know from recent space experiments and theoretical developments are manifestation of the interaction of solar plasma with the earth's magnetic field. Much before these ideas were established Alfvén knew the key to the problem of understanding the auroral phenomena and in his very early years of research began to study the motion of charged particles in electric and magnetic fields. The guided centre approximation given by Alfvén to study the equations of motion of charged particles in electromagnetic fields led to the development of this subject and the important adiabatic theory for describing the mathematical models of plasma based on single charged particle motion. The importance of Alfvén work in this area can hardly be exaggerated.

Alfvén developed several other concepts regarding interaction of plasmas with magnetic and electric fields. Some of these were not well received initially but have later proved to be highly relevant and useful. For example, in a visionary paper published in 1958, he proposed auroral primary electrostatic potential drop along geomagnetic field lines. Such electric fields with huge potential drops were considered impossible at that time, but modern space probe measurements have established such electric fields and have also shown that this mechanism indeed plays a crucial role in the physics of aurora. Another example: as a part of a theory of the origin of the solar system, Alfvén introduced the daring hypothesis of 'critical velocity' in plasma–neutral gas interaction. Its experimental validation came years later, and it took many years more before the mechanism could be explained theoretically.

Carl G. Fälthammar, Alfvén's student and close associate mentioned that though it is not widely recognized that as early as 1937 Alfvén argued for the existence of a galactic magnetic field and by inference an ionized interstellar medium, both of which were at that time considered impossible but now, of course, are well established

As to the theory of the origin of the Universe, Alfvén had his favourite slogan: 'In the beginning there was plasma'. To understand the electricity and magnetism in space and the interaction of electromagnetic fields with conducting ionized matter in the plasma Universe was his passion which did not leave him till the very last days of his life.

Hannes Alfvén contributed towards the progress of space and astrophysical sciences not only through his work but by inspiring many students, colleagues and friends all over the world. I pay my homage to this great scientist and will miss his inspiring correspondence.

CHANCHAL UBEROI

Department of Mathematics
Indian Institute of Science
Bangalore 560 012, India

Professor S. N. Bose's contribution to quantum statistics – A critical review

S. K. Das and S. Sengupta

'The Indian Bose has given a beautiful derivation of Planck's law including the constant (i.e. $8\pi\nu^2/c^3$)', wrote Einstein in July 1924 in a letter to Ehrenfest. He further commented that 'Bose's derivation is elegant but the essence remains obscure'[1]. We see that Bose's work, from the very beginning, evoked a sense of mystery which haunted physicists even after a lapse of more than 50 years. The first surprise is that quantum statistics, the last major discovery in quantum theory prior to the advent of quantum mechanics, was initiated by a man almost unknown to the world of physics, living in isolation in Dacca (now in Bangladesh), far removed from the main centres of hectic activities in this field. It is doubtful if there is a second example like this in the history of physics. Perhaps it is this unbelievable element in Bose's discovery which prompted the Nobel laureate scientist Max Delbruck to ask 'Was the Bose–Einstein statistics arrived at by serendipity?'[2]. The same spirit is revealed in Pais' remark that 'Bose's work is the most successful shot in the dark'[3]. There is, however, a more important reason to suspect that the discovery might have been accidental. The first three papers on the subject published in quick succession (the first two papers of Bose[4,5] published in July 1924 and the first paper of Einstein[6] in September 1924) make no mention of the most significant step forward in quantum statistics, viz. the concept of indistinguishability of the particles. It is really surprising that neither Bose nor Einstein in their first paper considered it necessary to examine critically the derivation of the new expression for thermodynamic probability. It is only in his second paper[7] (published in December 1924) that Einstein discussed this problem in detail after Ehrenfest and others pointed out the gap.

In this connection we mention a few more interesting features not quite well known. Bose actually gave two different expressions for thermodynamic probability in his two papers. In the first paper[4] he writes

$$W_1^s = \frac{Z^s!}{p_0^s!\, p_1^s! \cdots} \quad W = \Pi_s W_1^s \qquad \text{(B1)}$$

where Z^s is the total number of elementary phase cells belonging to photons of energy $h\nu^s$, in which N^s photons are distributed. p_r^s is the number of cells which contain r photons. As a result we have

$$\sum_r p_r^s = Z^s \quad \text{and} \quad \sum_r r p_r^s = N^s.$$

In the second paper, Bose writes

$$W_2^s = \frac{(Z^s + N^s - 1)!}{N^s!\,(Z^s - 1)!} \approx \frac{(Z^s + N^s)!}{N^s!\, Z^s!}$$

and $W = \Pi_s W_2^s$. (B2)

Before giving the expression (B1), Bose simply writes 'now it is easy to calculate the thermodynamic probability (macroscopically defined) of a state'. He gave no idea as to how he arrived at the expression or why he chose the particular description of the macroscopic state. In the second paper, while discussing the thermodynamic probability for radiation, Bose writes that 'this has been derived earlier'. And he gives two references. One is a paper by Bose to be published in *Philosophical Magazine*, but actually it was never published. The other is a paper by Debye (the paper published in 1910, which we shall discuss later). Thus, we do not have any record of the details of Bose's derivation of either (B1) or (B2).

Mehra[8], while commenting on the relation between the two expressions, says that W_1^s can be reduced to W_2^s. This, however, is wrong. W_1^s summed over all possible distributions of p_r^s leads to W_2^s. Similar imprecise remarks have also been made by Ghose[9] while comparing (B1) with (B2). Einstein used (B1) in his first paper and derived (B2) using the concept of indistinguishability, in the second paper. Pais[10] remarked wrongly that (B2) is the 'Einstein's expression for W now used in textbooks'. It is generally believed that Bose derived the photon statistics by dropping the constraint on N, the total number of photons. He kept only E, the energy, as constant. This is not quite correct. Such confusions arise because the differences in Bose's two derivations have never been examined critically. In his first paper, it appears that Bose maximized $W = \Pi_s W_1^s$, but in fact he was maximizing W_1^s. Here he was maximizing with respect to variables p_r^s (which are regarded as the macroscopic parameters here) and it turns out that the procedure is equivalent to maximizing each W_1^s separately keeping Z^s and $E^s = N^s h\nu^s$ constant. This implies that the photon number N^s is also kept constant. Pais[11] comments that 'the final irony is that the constraint of fixed Z^s is irrelevant'. This is a careless remark. In the logarithmic derivative of W_1^s, one has to use this condition. However, the additional input of this condition through Lagrange's multiplier turns out to be trivial and unnecessary. But this is only an accidental result. In his second paper, Bose maximized W_2^s with respect to the variables N^s keeping only $E = \Sigma_s N^s h\nu^s$ constant. Many such misconceptions seem to arise because physicists, in general, completely overlook Bose's second paper.

A more significant misconception is that quantum statistics started with Bose's work. In his second paper, Bose starts with a critical examination of a paper by Debye[12] published in 1910, where he used statistical mechanics to derive Planck's law. Bose concludes that Debye's derivation is consistent with the photon description except for his derivation of the factor $8\pi\nu^2/c^3$. Debye used an expression for W^s first written down by Planck in his book *Warmestrahlung* (English translation[13] appeared in 1913). This expression is identical to (B2) but Z^s and N^s were given a slightly different meaning. Natanson[14], in 1911, derived Planck's law using an expression for W^s which is identical to (B1). Both Debye and Natanson described black-body radiation in terms of normal-mode vibrations. This is equivalent to the description of radiation as an ensemble of harmonic oscillators. In 1914, Ehrenfest and Kamerlingh Onnes[15] published a paper in which they tried to interpret Planck's formula (B2) in terms of the distribution of photons. They concluded that if photons are regarded as mutually independent, Planck's formula cannot be obtained. (See refs. 8 and 9 for a brief summary of these and other works.)

The influence of these earlier works in the development of quantum statistics has mostly been ignored. The central mystery that Bose discovered the photon statistics without discovering the basic concept of indistinguishability of photons remains unresolved. Authors reporting Bose's work fill up the gap by asserting that implicitly Bose introduced the concept of indistinguishability. Such statements are not true and do not put Bose's work in the right perspective.

We suggest an alternative angle to look at Bose's contribution. Based on available facts and some plausible conjectures, we show that Bose was partly right when he said that he was doing a Boltzmann type of calculation. In this work he was strongly influenced by earlier works, particularly Planck and Debye. Thus, a continuity in the development of quantum statistics can be established. This is possible because of a fortunate circumstance. The concept of indistinguishability can be pushed back till we come to discuss the Bose–Einstein statistics of classical particles. For photons this concept is not an essential requirement. For example, if one uses Gibbs canonical ensemble formalism, one can avoid the concept of indistinguishability.

There is another interesting aspect of Bose's work. A photon when regarded as a particle has an unusual nonclassical property. It has only one energy state of energy $h\nu$ and momentum $h\nu/c$. If we try to change the momentum, we have to change ν and this will give us a new photon and not an excited state of the earlier photon. Thus, for an ensemble of N identical photons of energy $h\nu$, if $E = Nh\nu$ is kept constant, N becomes a constant and no variation is possible. Photons with different frequencies are not indistinguishable. If we consider radiation in an enclosure to consist of different frequencies ν^s, then through absorption and re-emission by a speck of matter the number N^s of each type of photons can be changed and variation of the probability function W becomes possible. But such an ensemble consists both of distinguishable and indistinguishable photons. This naturally raises the question, what type of statistical ensemble of photons was considered by Bose? Can this ensemble be compared with the standard Boltzmann or Gibbs ensembles used for classical particles? We discuss this question later.

Bose's two expressions for W: A conjecture on how Bose arrived at them

The basic difference between the two expressions for W which Bose used is in the definition of the thermodynamic macrostate. In (B1) this is defined by the variables p_r^s and in (B2) by N^s. The necessity for this arises because the ensembles considered are really different, though apparently they seem to be the same. We shall take up this question later. At present we are interested in the mathematical relation between the two expressions. This relation is rather complex and many misstatements are found in the literature. Let us first write the correct relations:

$$W_2^s = \sum_{(\cdots p_r^s \cdots)} W_1^s \quad . \qquad \text{(R1)}$$

Here s is kept fixed and the summation on the right-hand side is over all possible sets of distributions defined by $(\cdots p_r^s \cdots)$ keeping N^s and Z^s fixed. It is evident that $\Pi_s W_1^s \neq \Pi_s W_2^s$ for arbitrary distributions. Let us now suppose that we allow permutations among the photons, regarding them as classical particles. Then we get

$$\sum_{(\cdots p_r^s \cdots)} W_1^s \times \frac{N^s!}{\Pi_r (r!)^{p_r^s}} = (Z^s)^{N^s}. \qquad \text{(R2)}$$

The right-hand side gives the well-known classical value. The correctness of these relations can be checked only by showing that they treat the same combinatorial problem using different procedures. As we see, the relations are far more complex to be comprehended intuitively. It is a good and worthwhile exercise to take up a special case with small values of Z^s and N^s and verify the relations.

In the Planck–Debye derivations of (B2), there is a change of language. They consider radiation as an ensemble of quantized harmonic oscillators. Z^s is interpreted as the number of oscillators of frequency between ν^s and $\nu^s + d\nu^s$. Excitation energy of an oscillator is written as $(n^s + \frac{1}{2})h\nu^s$ and, of all the oscillators in Z^s, $\Sigma n^s h\nu^s$ + constant is the zero-point energy. Dropping the constant part, the total excitation energy of the oscillators is $E^s = N^s h\nu^s$, where $N^s = \Sigma n^s \cdot N^s$ may be defined as the total excitation number for the Z^s oscillators. The difference in the interpretations of the two quantities Z^s and N^s by Bose and Planck–Debye is very significant (Table 1).

It is to be noted that the mathematical expressions for Z^s and E^s are the same in the two interpretations. If we look at the combinatorial problem of distributing the number N^s in Z^s oscillators (Planck–Debye picture) then the result is the same as in (B2). Here Z^s-oscillators are all distinguishable and the question of identity of photons does not arise. Hence, the calculation follows a straightforward Boltzmann method. When we try to evaluate W^s using the Bose picture, two different results are obtained. If we drop permutation among photons, we get (B2). But if we allow permutation, we get $W^s = (Z^s)^{N^s}$. The central problem is that Bose, who was well conversant with Boltzmann's method of calculation, should have written down this expression. Instead he wrote (B2). Why?

In his statement to Mehra[16], Bose stated that he had read Planck's book *Warmestrahlung*. In his second paper[5] Bose examined Debye's paper critically. It is clear that he was familiar with the expression (B2) for W^s. 'Debye has shown that Planck's law can be derived using statistical mechanics', with these words Bose starts his second paper. But Bose

Table 1. Difference in the interpretations of Bose and Planck–Debye

	Bose	Planck–Debye
$Z^s = 8\pi\nu^{s^2} d\nu^s v/h^3$	Number of elementary phase cells, each of volume h^3, belonging to photons of frequency lying in the range ν^s to $\nu^s + d\nu^s$	Number of different harmonic oscillators belonging to the frequency range ν^s to $\nu^s + d\nu^s$
N^s	Total number of photons in Z^s cells	Total excitation number of the Z^s oscillators
$E^s = N^s h\nu^s$	Total energy of all the photons in Z^s	Total energy of all the oscillators in Z^s

objects that his derivation is not completely independent of classical electrodynamics. Because Debye calculates the number Z^s by using the idea of normal-mode vibrations, i.e. using the idea that radiation has a wave nature. Bose strongly believed in Einstein's photon picture of radiation. His one obsession was that if this idea is correct then one must be able to derive Planck's law by considering radiation as an ensemble of particles (photons) and in no way using the wave picture. Debye's use of normal modes in deriving Z^s was considered by Bose as a serious flaw in this scheme. So he replaced this by a new derivation by introducing the concept of phase space of photons each of which has a momentum $h\nu^s/c$. In this new picture Bose writes '$8\pi\nu^2\, d\nu\, v/c^3$ can be interpreted as the number of elementary cells in the six-dimensional phase space of the quanta. *Further calculations (of Debye) remain essentially unchanged*'. The last comment has a crucial significance. It means that Bose assumes that in the photon picture, Debye's calculation of W^s (i.e. (B2)) remains unchanged. By implication it follows that Debye's idea of an oscillator excited to a state of energy $n^s h\nu^s$ is, in Bose's language, equivalent to having n^s phonons each of energy $h\nu^s$. This equivalence is basically right as has been proved after quantization of radiation field. But in 1924 it was not known. Here Bose showed superb intuition. With this equivalence, both the expressions for Z^s and E^s become identical in the two pictures of radiation (ensemble of photons and ensemble of normal-mode vibrations). Hence, if Planck's law is to be derived, the expression for W^s must necessarily be the same in the two pictures. Our conjecture, based on the brief statements of Bose quoted above, is that Bose arrived at this conclusion through intuition and did not wait for a proper analysis of the implications. This conjecture alone can explain the way he introduced the expression for W^s for photons in his second paper – 'Thermodynamic probability for radiation: This has been derived earlier[7]'. Reference 7, as we have mentioned earlier, contains Debye's paper. Our conjecture also explains why Bose thought that he was not introducing any basically new concept in the calculation of W^s. Once the equivalence is taken for granted, one can utilize the Planck–Debye derivation of W^s, which is done entirely in the Boltzmann framework.

What about Bose's formula for W^s in the first paper? We do not know if Bose had seen Natanson's paper of 1911, where the expression (B1) was derived using the Planck–Debye picture. Naturally, here our conjecture will be more speculative. Our guess is that Bose was not aware of Natanson's paper. Otherwise, he would have referred to it as he had referred to Debye in his second paper. Planck, in his book, considered radiation of a single frequency. His ensemble consisted of N^s quantum oscillators all of frequency ν^s. For this he derived the expression for W^s (i.e. (B2)). But here both the parameters Z^s and N^s are constant and there is no variable parameter with respect to which one can maximize W^s. Both Natanson and Bose independently resolved this difficulty by introducing the variable p_r^s. Bose, using photon language, had the additional constraint to make his expression for W_1^s consistent with the Planck's expression for W_2^s.

After enunciating the combinatorial problem as Bose had done in the first paper, one would immediately write down the expression for W_1^s. The question of permutation of the particles would come later. Our conjecture is that after writing W_1^s, Bose could easily check that it was consistent with W_2^s. So he was sure that the formula was correct, and he overlooked the finer details about the permutation of the particles. One thing seems almost sure, that Bose was not aware of the paper by Ehrenfest and Kamerling Onnes[15]. Had he seen this paper, it seems almost certain that the question of indistinguishability would have taken a prominent part in his paper.

The scenario changed completely when Einstein extended Bose's formula to classical particles. The alternative Planck–Debye picture is no longer valid and one has to justify the derivation of W_1^s or W_2^s independently. Hence, the concept of indistinguishability becomes essential.

Let us now summarize our conclusions. Firstly, Bose reinterpreted the Planck–Debye derivation of radiation formula in terms of photon language. In doing this he correctly anticipated that both the languages are true and that the expression for W^s remains unchanged by the change of language. This equivalence made it possible for Bose to derive photon statistics without introducing the concept of indistinguishability. Secondly, Bose introduced the concept of phase space of photons, which made his photon statistics look almost identical to particle statistics. This led to the extension of his formula for W^s to classical particles. In the Planck–Debye language one could not even think of such an extension.

Statistical ensemble in Bose's two derivations

We have already mentioned some very peculiar nonclassical properties of photons. As a result, an ensemble of photons will have properties not exactly similar to those of an ensemble of classical particles. Hence, it is relevant to discuss precisely what sort of ensembles Bose used in his photon statistics.

In the first paper Bose writes $S = k \ln W$, $W = \Pi_s W^s$ (W^s given by (B1)). The constraints are $Z^s = \Sigma_r p_r^s$, $E = \Sigma_s \Sigma_r h\nu^s r p_r^s$. Apparently, it may seem that Bose is considering an ensemble of photons of all frequencies and maximizing W. But, in fact, he was maximizing only W^s keeping Z^s and $E_s = N^s h\nu^s$ constant. It is easy to see that, if we do this, all of Bose's results remain the same. This is because each ensemble of N^s photons has a most probable distribution for p_r^s corresponding to the equilibrium state at a temperature T, say. If all these ensembles are put in thermal contract, they will still be in equilibrium. Thus, maximizing W^s in terms of p_r^s automatically maximizes W.

Thus, Bose, in the first place, was considering a Boltzmann ensemble of N^s identical photons each of one energy stage $h\nu^s$. The nonclassical feature of this ensemble is that when E^s is kept fixed, no variation of W^s is possible. The situation is somewhat like an ensemble of particles all of whose energy states are degenerate. Bose introduced a new definition of a macroscopic state in terms of the parameters p_r^s, which can then be varied to give different values of W^s. It is interesting to note that in terms of the Planck–Debye picture, variables p_r^s come out quite naturally. The ensemble of Z^s oscillators with a total excitation energy $E^s = N^s h\nu^s$ is a simple Boltzmann ensemble, where the distribution parameter is the number of oscillators excited to the rth state of energy $rh\nu^s$, and this exactly corresponds to the variable p_r^s

defined by Bose. This again confirms our conjecture that Bose was intuitively convinced about the mathematical identity of the two descriptions of radiation.

In the second paper, Bose, following Debye, maximizes W and not W^s. Here the distribution parameter is N^s. The ensemble is neither Gibbs canonical nor Boltzmannian. It is semi-Gibbsian. We have N^s identical photons as a subensemble. Varying ν^s, we generate other subensembles and put them in thermal contact. Photons in any two subensembles are nonidentical. Since W^s depends on Z^s and N^s, we can change W^s by varying N^s and then maximize W by the usual technique. The only constraint here is that $E = \Sigma_s N^s h\nu^s$, a constant.

Summary

Bose had been plagued throughout his life by the question how he got the formula for W^s without using the indistinguishability concept. In an interview with Mehra[17] towards the end of his life, he stated: 'I had no idea that what I had done was really novel . . . Instead of thinking of the light quantum just as particles, I talked about these states. Somehow, this was the same question which Einstein asked when I met him (in October or November 1925), how had I arrived at this method of deriving Planck's formula.' Bose thought that he was using the standard Boltzmann method. And we believe that he was right.

The development of quantum statistics is not as abrupt as is usually believed. There is a continuity starting from the works of Planck, Debye, Natanson, Ehrenfest and Kamerling Onnes and others. This has been possible because, for radiation, there are two alternative ways of treatment; one may consider it as an ensemble of distinguishable quantized harmonic oscillators (Planck–Debye picture) or as an ensemble of indistinguishable photons (Bose picture). Both lead to the same expression for W. Bose used the photon language to describe the ensemble but the Planck–Debye picture to calculate W and thereby sidetracked the question of indistinguishability. When Einstein extended Bose's expression for W to particles, then the alternative Boltzmann picture of oscillators disappeared. One can derive the particular expression for W by using the concept of indistinguishability alone. And this Einstein did in his second paper.

In the history of physics, the discovery of the B–E statistics is a remarkable example to show how tortuous the path along which ideas in physics develop is. We are reminded of the incisive remark by Koestler[18], made in connection with the discoveries in cosmology: 'The history of cosmic theories may without exaggeration be called a history of collective obsessions and controlled schizophrenia; and the manner in which some of the most important individual discoveries were arrived at reminds one more of a sleep-walker's performance than an electronic brain's.'

1. Einstein, A., Letter to P. Ehrenfest, July 12, 1924.
2. Delbruck, M., *J. Chem. Ed.*, 1980, **57**, 467–470.
3. Pais, A., *Subtle is the Lord*, Oxford University Press, 1982, p. 428.
4. Bose, S. N., *Z. Phys.*, 1924, **26**, 178–181. (English translation in *Physics News*, 1974, **5**, 40–42).
5. Bose, S. N., *Z. Phys.*, 1924, **27**, 384–393. (English translation in *Physics News*, 1974, **5**, 42–46).
6. Einstein, A., *Preuss Akad. Wiss., Phys. Math.Klasse, Berlin*, 1924, **XXII**, 261–267.
7. English translation of ref. 6 in *Physics News*, 1983, **14**, 36–41.
8. Mehra, J. and Rechenberg, H., *The History of Development of Quantum Theory*, Springer, 1982, vol. 1, Part 2, pp. 554–578.
9. Ghose, P., *S. N. Bose, The Man and His Works*, Part I, S. N. Bose National Centre for Basic Sciences, 1994, p. 35.
10. Pais, A., *Subtle is the Lord*, Oxford University Press, 1982, p. 431.
11. Pais, A., *Subtle is the Lord*, Oxford University Press, 1982, p. 427.
12. Debye, P., *Ann. Phys.*, 1910, **33**, 1427–1434.
13. Planck, M., *The Theory of Heat Radiation*, Dover, 1913, p. 146.
14. Natanson, L., *Phys. Z.*, 1911, **12**, 659–666.
15. Ehrenfest, P. and Kamerling Onnes, H., *Proc. Kon. Akad. Wetensch (Amsterdam)*, 1914, **17**, 870–873.
16. Mehra, J. and Rechenberg, H., *The History of Development of Quantum Theory*, Springer, 1982, vol. 1, part 2, p. 564.
17. Mehra, J., *Biogr. Mem. Fellows R. Soc. (London)*, 1975, **21**, 117–154.
18. Koestler, A., *The Sleep-Walkers*, Penguin, 1964.

S. K. Das is presently in the Department of Physics, Krishnagar Govt. College, Krishnagar, Nadia, West Bengal 741 101, India and S. Sengupta is in Condensed Matter Physics Research Centre, Department of Physics, Jadavpur University, Jadavpur, Calcutta 700 032, India.

Edited and published by Prof. S. Ramaseshan and Prof. P. Balaram, Current Science Association, Bangalore 560 080.
Typeset by R D Typesetters, Bangalore 560 010. Printed by Lotus Printers, Bangalore (Ph: 3380167).

MEETINGS/SYMPOSIA/SEMINARS

32nd Annual Convention of Indian Geophysical Union and Meeting on Marine Geophysics

Place: National Geophysical Research Institute, Hyderabad
Date: 19–21 December 1995

Topics include: 1. Geophysical instrumentation. 2. Atmospheric sciences. 3. Geophysical exploration. 4. Space sciences and planetology. 5. Geodynamics. 6. Lithospheric studies. 7. Marine geosciences.

Contact: Dr P. R. Reddy
Hon. Secretary, Indian Geophysical Union
NGRI Campus, Hyderabad 500 007
Phone: 670141

International Symposium on Geology and Geophysics of the Indian Ocean

Place: National Institute of Oceanography, Goa, India
Date: 21–24 October 1996

Topics: 1. Structure and tectonics. 2. Particle fluxes and sedimentary processes. 3. Sea-level changes and paleoceanography. 4. Non-living resources of continental margins. 5. Deep-sea mineral deposits and their environment. 6. Technology, policy, economics and geopolitics of offshore exploration and exploitation

Contact: Dr Ch. M. Rao, Convener
National Institute of Oceanography
Dona Paula, Goa 403 004, India
Phone: 91-(0)832-221352
Fax: 91-(0)832-223340/91-(0)832-221360 (BTIS)
e.mail: ocean@bcgoa.ernet.in

4th International Conference on Physiological Fluid Dynamics

Place: Jiwaji University, Gwalior, India
Date: 8–10 December 1995

Topics include: Biorheology, Atherosclerosis, Flow visualization, Clinical hemorheology, Simulation/modelling, Biotransport, Microcirculation, Arterial hemodynamics, Red cell deformability and aggregation.

Contact: Prof. V. P. Saxena, Secretary
School of Mathematics & Allied Sciences
Jiwaji University, Gwalior 474 011, India
Phone: (0091) (751) 341905
Fax: (0091) (751) 341450

Indian Association of Special Libraries & Information Centres (IASLIC) XX All India Conference on Information Technology Products of 1990s and Library and Information Centres

Place: Dept. of Library & Information Science
Lucknow University, Lucknow, U.P.
Date: 26–29 December 1995

Topics include: Information technology as media, Computer software information products, Information seeking behaviour and information technology products, Infrastructural facilities needed for utilisation and maintenance of information products, Economic feasibility studies of information technology products, Legal implication to utilisation of information technology products, Competency in utilisation of information technology products knowledge and skills.

Contact: Convener
Editorial Committee
XX All India Conference
IASLIC, P-291 CIT Scheme 6M, Kankurgachi
Calcutta 700 054
Phone: 349651

National Symposium on Recent Advances in Biosciences

Place: Dept. of Biosciences, Maharshi Dayanand University, Rohtak
Date: 14–16 October 1995

Contact: Prof. V. P. Agrawal
Secretary General
Society of Biosciences, 25/4, Ram Bagh Road
Musaffarnagar 251 001, U.P.

National Symposium on Deccan Flood Basalts of India

Place: Nagpur
Date: February 1996

Topics include: 1. Stratigraphy, structure and tectonics. Geophysical studies leading to understanding of sub-Deccan crust and mantle. 2. Geological and volcanological studies. Morphological characters of lava flows and interpretation of volcanic history. 3. Petromineralogical, geochemical and petrogenetic studies. 4. Genesis and evolution of magma, composition of sub-continental mantle, degree of crustal contamination and genesis of small volumes of magmas of extreme compositions. 5. Age and duration, palaeomagnetism, palaeontological studies on infra and intertrappean beds. 6. Geomorphology, plateau uplift and erosion, recognition and dating of planation surfaces. 7. Intrusives: Geology, geochemistry, structure and tectonics, age, palaeomagnetic characters, spatial and temporal relationship between intrusives following different tectonic trends and of differing compositions, recognition of feeder dykes. 8. Geophysical, geotechnical and seismic studies. 9. Economic aspects, including groundwater potential, with special emphasis on geothermal energy resources and prospects for polished stone industry.

Contact: Shri K. K. K. Nair
Geologist (Sr), Organising Secretary
Geological Survey of India
Central Region, New Complex, Seminary Hills
Nagpur 400 006
Phone: (O) 534448 (R) 220089
Fax: (0712) 532636

CURRENT SCIENCE

A fortnightly journal of research

Editors: S. Ramaseshan and P. Balaram

Editorial Board

S. Arunachalam, Central Electrochemical Research Institute, Karaikudi 623 006
Asis Datta, School of Life Sciences, Jawaharlal Nehru University, New Delhi 110 067
D. Balasubramanian, Centre for Cellular and Molecular Biology, Hyderabad 500 007
J. Chandrasekhar, Department of Organic Chemistry, Indian Institute of Science, Bangalore 560 012
R. Gadagkar, Centre for Ecological Sciences, Indian Institute of Science, Bangalore 560 012
K. N. Ganeshaiah, Department of Plant Genetics and Breeding, University of Agricultural Sciences, Bangalore 560 065
R. V. Hosur, Tata Institute of Fundamental Research, Bombay 400 005
N. V. Joshi, Centre for Ecological Sciences, Indian Institute of Science, Bangalore 560 012
Jyotirmoy Das, Indian Institute of Chemical Biology, Calcutta 700 032
C. C. Kartha, Division of Cellular and Molecular Cardiology, Sree Chitra Tirunal Institute for Medical Sciences and Technology, Thiruvananthapuram 695 011
S. S. Krishnamurthy, Inorganic and Physical Chemistry Department, Indian Institute of Science, Bangalore 560 012
S. C. Lakhotia, Department of Zoology, Banaras Hindu University, Varanasi 221 005
K. Muralidhar, Department of Zoology, Delhi University, Delhi 110 007
R. Narasimhan, CMC Limited, Bangalore 560 001
K. Neelakantan, ANURAG (Advanced Numerical Research and Analysis Group), Hyderabad 500 258
R. Nityananda, Raman Research Institute, Bangalore 560 080
T. Padmanabhan, Inter-University Centre for Astronomy and Astrophysics, Pune 411 007
B. R. Pai, Propulsion Division, National Aerospace Laboratories, Bangalore 560 017
T. J. Pandian, School of Biological Sciences, Madurai Kamaraj University, Madurai 625 021
Y. S. Rajan, Department of Science and Technology, New Delhi 110 016
M. Ramakrishnan, Geological Survey of India, Calcutta 700 016
A. V. Ramani, T. T. K. Pharma Ltd., Bangalore 560 025
S. Ramaswamy, Department of Physics, Indian Institute of Science, Bangalore 560 012
A. K. Raychaudhuri, Department of Physics, Indian Institute of Science, Bangalore 560 012
S. R. Shetye, Physical Oceanography Division, National Institute of Oceanography, Dona Paula, Goa 403 004
V. Siddhartha, 51 Bharati Nagar, New Delhi 110 003
V. Srinivas, School of Mathematics, Tata Institute of Fundamental Research, Bombay 400 005
G. Srinivasan, Raman Research Institute, Bangalore 560 080
R. Uma Shaanker, Department of Crop Physiology, University of Agricultural Sciences, Bangalore 560 065
K. S. Valdiya, Department of Geology, Kumaun University, Nainital 263 002
G. Venkataraman, Department of Physics, Sri Sathya Sai Institute of Higher Learning, Prasanthi Nilayam 515 134
M. Vidyasagar, Centre for Artificial Intelligence and Robotics, Bangalore 560 001
K. VijayRaghavan, TIFR Centre, Indian Institute of Science Campus, Bangalore 560 012

Current Science, founded in 1932, is published by the Current Science Association in collaboration with the Indian Academy of Sciences. The journal is also intended as a medium for communication and discussion of important issues that concern science and scientific activity. All articles published in *Current Science*, especially editorials, opinions and commentaries, letters and book reviews, are deemed to reflect the individual views of the authors and not the official points of view, either of the Current Science Association or of the Indian Academy of Sciences.

Working Committee, Current Science Association

P. Balaram, Indian Institute of Science, Bangalore
R. Chidambaram, Department of Atomic Energy, Bombay
V. K. Gaur, CSIR Centre for Mathematical Modelling and Computer Simulation, Bangalore
K. Kasturirangan, ISRO Satellite Centre, Bangalore
P. K. Kaw, Institute for Plasma Research, Gandhinagar
N. Kumar, Raman Research Institute, Bangalore (*Vice-President*)
N. V. Madhusudana, Raman Research Institute, Bangalore (*Secretary*)
A. Mani, Field Research Unit, Indian Institute of Tropical Meteorology, Bangalore (*Past-President*)
R. A. Mashelkar, National Chemical Laboratory, Pune
R. Narasimha, Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore
G. Padmanaban, Indian Institute of Science, Bangalore (*Vice-President*)
R. Ramachandran, The Institute of Mathematical Sciences, Madras
P. Rama Rao, Department of Science and Technology, New Delhi
S. Ramaseshan, Raman Research Institute, Bangalore (*Treasurer*)
M. S. Valiathan, Manipal Academy of Higher Education, Manipal (*President*)

Editorial Office

Editorial staff

Chandrika Ramesh, G. Madhavan, C. S. Ravi Kumar, M. S. Venugopal

Circulation and Accounts

Peter Jayaraj, Ranjini Mohan, B. Sethumani, Shanthi Bhaskar, B. K. Shivaramiah

Annual subscription 1995

	Institutional	Personal
India	Rs 350	Rs 100
All other countries (AIR MAIL)	US $175	US $50

Single copies other than special issues: Rs 35/US $15

Editorial Office: CURRENT SCIENCE
C. V. Raman Avenue
P.B. 8001
Bangalore 560 080, India
Telephone: 91-80-3342310
Telex 845-2178 ACAD IN
Fax: 91-80-3346094

Copyright © 1995, Indian Academy of Sciences

INFORMATION FOR CONTRIBUTORS

GENERAL

All manuscripts should be addressed to the Editor, *Current Science*, P. B. No. 8001, C. V. Raman Avenue, Bangalore 560 080. Submission of an article will be held to imply that it has not been previously published and is not under consideration for publication elsewhere; and further, that if accepted, it will not be published elsewhere. *Three copies of contributions of all categories* are required, with a letter of transmittal giving (i) names and complete addresses of the authors and (ii) title of the contribution and the category in which it is submitted (see below).

Current Science is a multidisciplinary journal and therefore research and review papers of general significance that are written clearly and well organized will be given preference. All papers will be first assessed by a Reviewing Editor. Papers found unsuitable in terms of the overall requirements of the journal will be returned to the authors. The others will be sent for detailed review. *Both solicited and unsolicited material will be reviewed.* Authors of these papers will be notified of acceptance, rejection, or need for revision of the paper. Returned papers cannot be resubmitted.

Illustrations and other material to be reproduced from other publications must be properly credited; it is the authors' responsibility to obtain permission for reproduction (copies of letters of permission should be sent).

CATEGORIES OF MANUSCRIPT

General articles (not exceeding 5000 words) discuss current trends in research in a field that will be of interest to readers outside the field; interdisciplinary topics: science policy and administration; or some aspect of the application of science and technology to human needs or the impact of science and technology on society/ecosystems/life. They should include a summary not exceeding 100 words, introductory paragraph(s), brief subheads at appropriate places to point to what follows, illustrations that will help a general reader, and references.

Review articles (not exceeding 5000 words) are expected to survey and discuss recent developments in a field. They should be well focused and organized, and avoid a general, 'textbook' style.

Research articles (not exceeding 4000 words) should report research results of fairly major significance. They should include an abstract not exceeding 100 words, introductory paragraph(s), and brief subheads.

Research communications (not exceeding 2000 words) should contain important findings that are novel and of fairly broad interest. They should include a brief abstract and an introductory paragraph. Text should not be broken up under subheads.

Correspondence includes letters, not exceeding 500 words, that are of general interest to scientists. All letters cannot be published.

Scientific correspondence contains technical comments, including those on articles or communications published in *Current Science* within the previous six months. Letters may be reviewed and edited.

Research news articles are intended to inform nonspecialists about recently published advances or important findings discussed at a meeting. Authors should also send a copy of the paper(s) on which the article is based. Meeting reports should avoid merely listing brief accounts of topics discussed, and must convey to readers the significance of an important advance.

Research accounts articles are intended to be personalized reviews of research from the authors' own laboratory, based on a body of published work. The articles must provide appropriate background to the area in a concise introduction, which should also serve to place the author's work in proper perspective. Articles will normally not exceed 8 to 10 printed pages.

Opinion articles present views on issues related to science and scientific activity. **Commentary** articles should contain expository notes on issues related to science and scientific activity.

Book reviews. Unsolicited reviews will also be considered. Reviews that merely 'list' brief descriptions of the contents cannot be published. Reviews should have 'context' and convey some information about the subject of the book.

Historical commentary and notes inform readers about interesting aspects of personalities or institutions of science or about watershed events in the history/development of science; most are expected to relate to India. Illustrations are welcome. Brief items will also be considered.

MANUSCRIPT PREPARATION

Manuscripts should be typed double-spaced on one side of white bond paper (21 × 28 cm). The pages should be numbered consecutively, starting with the title page and through the text, reference list, tables and figure legends. The **title** should be brief, specific and amenable to indexing. Not more than five **keywords** should be indicated separately; these should be chosen carefully and must not be phrases of several words. **Summary** and **abstract** should not have more than 100 words and should convey the main point of the paper, outline the results and conclusions, and explain the significance of the results.

Text. All papers should have a brief introduction. The text should be intelligible to readers in different disciplines and technical terms should be defined. Tables and figures should be referred to in numerical order. All **symbols** and **abbreviations** must be defined, and used only when absolutely necessary. Superscripts and subscripts and ambiguous characters should be clearly indicated. **Units of measure** should be metric or, preferably, SI. Methods should, as far as possible, be described briefly in appropriate table and figure legends.

Figures. In the case of line drawings one set of originals (without any lettering) is sufficient, with two copies containing lettering. In the case of photographs good prints are required with each copy of the manuscript; photocopies are not acceptable. Line drawings should be roughly twice the final printed size. The correct orientation should be indicated if not clear.

Photomicrographs and other photographs that require it must have a scale bar, which should be defined clearly in the legend. Primary data should be submitted as far as possible (e.g. actual photographs of electrophoretic gels rather than idealized diagrams).

References should be numbered in the order in which they appear, first through the text and then through table and figure legends. The following are examples of ways of writing references.

1. Mukundan, T. and Kishore, K., *Curr. Sci.*, 1991, **60**, 355–362.
2. Constantine, G., in *Biology of Bats* (ed. Wimsatt, W. A.), Academic Press, New York, 1970, vol. 1, pp. 319–322.

Acknowledgements should be brief. Footnotes are not allowed except to identify the corresponding author if not the first.

Cover photographs. Good photographs (colour or black and white) that pertain to a submitted paper will be considered for use on the cover. Good *prints* and a legend should be submitted with the manuscript. In the case of a colour picture, a transparency will be required for printing if accepted.

PROOFS AND PUBLICATION

Two sets of galley proofs are sent to the corresponding author. A reprint order form accompanies the proofs.

Founded: 1932

ISSN: 0011-3891

CURRENT SCIENCE

Volume 69 Number 2

25 July 1995

83 **In this issue**

CORRESPONDENCE

85 **Patenting biotechnologies**
Suresh Chandran

86 **Why not platinum?**
R. Shankar

87 **The Academy document: A perceptive appraisal of science education in India**
Atanu Kumar Pati

88 **The Academy document: A teacher's view**
D. P. Khandelwal

OPINION

89 **The Indian Academy of Sciences document on science education in universities: A criticism**
V. Sitaramam

SCIENTIFIC CORRESPONDENCE

94 **Development of cultured tomato anther to a fruit-like structure accompanied by *in vitro* ripening**
Pious Thomas and J. B. Mythili

96 **Genotoxicity of 'gudakhu', a tobacco preparation used in Orissa**
R. K. Das and B. C. Dash

97 ***Azolla* sp. from the Early Cretaceous of Cauvery basin, South India**
R. S. Patil, M. Rama Rao and C. G. K. Ramanujam

99 **Angioarchitecture of the glomerulus of a swamp catfish, *Chaca chaca* (Ham.) (Chacidae, Siluriformes)**
J. Ojha

GENERAL ARTICLE

101 **Location and magnitude of the 1853 Nepal earthquake and its relation to the rupture zones of contiguous great Himalayan earthquakes**
Roger Bilham

REVIEW ARTICLES

129 **How fast is ultrafast chemistry?**
Biman Bagchi

140 **Membrane current and potential change during nerotransmission in smooth muscle**
Rohit Manchanda

RESEARCH ARTICLES

151 **Flexicrystallography: Curved surfaces in chemical structures**
Alan L. Mackay

161 **Observed and theoretical acceleration response spectra in the Tehri region Implications for seismic hazard in the region**
K. N. Khattri

RESEARCH COMMUNICATIONS

171 **Sr isotopes in rivers of India and Pakistan: A reconnaissance study**
J. R. Trivedi, Kanchan Pande, S. Krishnaswami and M. M. Sarin

178 **The non-induction of HSP70 in heat shocked Malpighian tubules o *Drosophila* larvae is not due to constitutive presence of HSP70 or HSC70**
Bhupendra N. Singh and S. C. Lakhotia

182 **Codon analysis of cyanobacterial genes**
S. Krishnaswamy and S. Shanmugasundaram

185 **Effects of feeding mulberry (*Morus* sp.) leaves supplemented with differen nutrients to silkworm (*Bombyx mori* L.)**
A. A. Sarker, M. R. Haque, M. A. Rab and N. Absar

BOOK REVIEWS

189 **The Guide to the Galaxy, *reviewed by* V. Radhakrishnan**

189 **Crustal Evolution of Singhbhum–North Orissa, Eastern India, *reviewed b* Dhruba Mukhopadhyay**

191 **Annual Review of Phytopathology 1994, *reviewed by* Y. L. Nene**

PERSONAL NEWS

192 **Professor Willy Fowler (1911–1995) – an obituary**
J. V. Narlikar

GENERALIA

193 **Who should look at stars**
Ashok S. Ganguly

HISTORICAL COMMENTARY AND NOTES

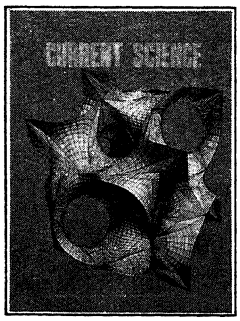
197 **Scientific tradition and other traditions**
B. M. Udgaonkar

Correction

128 **Is there dimorphism for style lengths in monoecious figs?**
Prarthana Kathuria, K. N. Ganeshaiah, R. Uma Shaanker and R. Vasudeva
[*Curr. Sci.*, 1995, **68**, 1047–1049]

Erratum

128 **The Great Bengal Cyclone of 1737 – An enquiry into the legend**
A. K. Sen Sarma [*Curr. Sci.*, 1995, **68**, 114–117]



COVER. An illustration of a curved surface expressed in terms hypergeometric functions. See page 151.

Indexed in CURRENT CONTENTS/GEOBASE/CHEMICAL ABSTRACTS

In this issue

High seismic hazard along the central Himalaya (Dehra Dun to Muzaffarpur), and risk to Tehri dam

The first step in estimating earthquake hazard in an area, from which engineering specifications could be derived for the design of earthquake-resistant structures, is to define the possible earthquake threat(s) to the region. This, in turn, requires (a) delineation of the most likely fracture plane(s) (geological faults) that could be the site where long-accumulated strain will be suddenly released and (b) rates of strain accumulation and the current strain budget(s) of these faults or fault systems, the last of these being difficult to determine directly except through stress measurements and/or the historical records of past earthquakes associated with the fault(s). Assuming a failure strain of 10^{-4}, one can then estimate the probabilities of exceedance of strain release or slip on these faults over a given period of time in the future, and further, as shown by Khattri (**page 161**) the expected ground motion time histories and the probabilities of exceedance of ground motion parameters (displacement, velocity, acceleration and duration of strong shaking) at a specified distance from them. The latter constitute the basic quantities needed for a seismic design.

Earthquake hazard in the Himalaya and in the adjoining densely populated plains to the south is dominated by the great plate boundary earthquakes which accompany a sudden slip of the persistently northward driving Indian plate, beneath the overriding Himalaya. Once every few hundred years, therefore, when the steadily accumulating strains exceed the bearing capacity of the intervening fault systems, various segments of the Himalaya lurch forward over the foothill plains in a giant leap of 6–10 m along a gently northward dipping fracture plane, which is typically 300 km long and 100 km wide.

However, the entire 2400 km long Himalayan belt from Nangaparbat in the west to Namcha Barwa in the east, although compressed at approximately the same rate, does not slip as a single continuous rupture. For, there are a number of transverse fractures incised by the strong preexisting grains of the underthrusting Indian plate, that segment the Himalayan arc mechanically, thereby providing an additional degree of freedom which allows its various segments to slip independently, adjusting the times of their slippage (earthquakes) according to the nature and strength of their respective fault asperities. Thus, the giant slips of the various Himalayan segments whilst having approximately equal renewal intervals (recurrence periods of great earthquakes) tend to stagger in time.

Four great earthquakes have been known to rupture about half of the Himalayan plate boundary over the past 100 years, and although a direct determination of the strain rate across this boundary still remains to be made, various constraints indirectly indicate a figure of 15 mm/yr (strain rate $\approx 1.5 \times 10^{-7}$/yr). This, in turn, suggests that the recurrence interval for a great Himalayan earthquake would be about 600 years. Unfortunately, our historical record of seismicity in India is rather poor, which makes it impossible to estimate the current strain budget along a fault that has no known record of an associated earthquake in the past, even if the rate of strain accumulation is known.

It will thus be apparent that whilst it would be safe to assume that those parts of the Himalaya which have ruptured in the past 100 years are unlikely to accumulate enough strain for a great earthquake at least for the next 400–500 years, others that have not been ruptured for 500–600 years by a great earthquake have most likely accumulated more than 8 m of slip (strain $\approx 10^{-4}$) and are, therefore, ripe for a great earthquake which could have a magnitude greater than 8.5.

One such long segment in the Himalaya where no great earthquake has occurred since the 13th century and where potential seismicity is, therefore, extremely high is the Central Himalayan seismic gap extending from Dehra Dun to Muzaffarpur. Since a number of large dams are being constructed in this area, an evaluation of its seismic potential over the next 300–400 years, covering the life and residual after-effects of the dam, has been a matter of great public concern and regrettably of an unresolved public debate between seismologists and engineers, particularly in respect of low estimates of earthquake hazard and high percentile of risk adopted in the engineering design of the dam structure.

Two well researched papers in this issue (Bilham, **page 101**) and (Khattri, **page 161**) address the two main arguments underlying the adoption of low hazard figures: (a) the assumption that the 1833 Nepal earthquake did considerably diminish the accumulated strain in this part of the Himalaya and (b) that the rock attenuation figures in the area being high, ground accelerations would decrease rapidly away from the fault. The two papers contained in this issue provide enough evidence to refute both these possibilities, and call for attention to the dangerous consequences of inadequate safety figures incorporated in the design of this large critical structure.

The first paper by Bilham shows convincingly that the only major earthquake to occur in this region after the well-documented disastrous earthquake of 1255 was the 1833

Nepal earthquake, which more or less coincided with the western extremity of the 1934 Bihar earthquake rupture and, in any case, was too moderate to cause any significant reduction of the strain budget in the Central Himalayan seismic gap. The message, therefore, is that this sector of the Himalaya has not experienced a great earthquake for about 750 years and unless accumulating strains are relieved by aseismic creep, which appears unlikely on the basis of the evidence of recent levelling data in Nepal, the accumulated slip in the region may be as high as 10 m, capable of generating a great earthquake of magnitude 8.5–9.0 in the near future.

The second paper by Khattri, using the actual ground acceleration records of the 1991 Uttarkashi earthquake as a basis for simulating the ground acceleration record at Tehri, which lies in the Central Himalayan seismic gap, shows that peak ground acceleration expected to occur at Tehri in the wake of $M = 8.5$ earthquake would be as high as 1.0 g, and would continue to shake the ground intensely a period longer than 20 s.

These two papers thus make a valuable and timely contribution towards alerting all those concerned with the construction of large critical facilities in the Central Himalaya to study carefully the scientific results and computer-simulated visualizations that argue against the unrealistically low figure of earthquake hazard adopted in the design of the 260 m high Tehri dam and to revise its design so as to ensure long-term protection to large unwary communities downstream.

V. K. G.

Molecular processes: Faster and faster

Chemical kinetics is the study of the rates of chemical reactions and their interpretation in terms of molecular mechanisms. Classically, the determination of rates was restricted to processes occurring in time scales of several seconds or longer due to limitations of measurement. Over the last forty years or so there has been a steady improvement in the shutter speed of chemists' cameras. While measurements in the millisecond regime were at the forefront in the 1950s, even picosecond time scales will be passé in the late 1990s. Technical barriers have been regularly breached and experimental studies of femtosecond (10^{-15} s) processes are now possible.

Part of the motivation for studying ultrafast molecular processes stems from the increasing inroads that have been made in understanding critical biological events in ultimate chemical detail. Electron transfer processes in photosynthesis and the critical isomerization of the visual pigment retinal in the rhodopsins following photoexcitation are among the best studied. It is after all enormously exciting to be able to follow events immediately after photon capture which eventually lead to the wonderful phenomenon of vision. Biology affords the physical scientist an extraordinary range of immensely complex problems, which may be addressed using techniques that have been tried and tested on the more tractable problems of physical chemistry.

Ions and molecules in solution are solvated and there is an unceasing dance of chemical species in solution, modulated by myriad interactions. Excitation of chromophores most often leads to changes in charge distribution, necessitating solvent rearrangements to achieve an equilibrium excited state. Measurements of solvent reorganization times, sometimes of the order of 200 fs to 1 ps, have become possible recently using ultrashort laser pulses, providing remarkable temporal resolution of molecular processes. The detection of ultrafast events in liquids like water, methanol and acetonitrile, of course, raises the question of their origin. This is happy hunting ground for theoreticians and provides many insights into the dynamic behaviour of liquids.

For the connoisseur of ultrafast events, electron transfer is undoubtedly one of the most delectable items on the menu of chemical processes. The celebrated Marcus theory is the cornerstone in the understanding of electron transfer. Can solvents or the medium affect the very rapid electron transfer reaction? Recent studies suggest that there are specific situations where medium effects are indeed observable. What appears to be a rather esoteric area of physical chemistry may acquire broad importance in the light of the determination of the detailed molecular structures of proteins involved in biological electron transfer, where electrons hop between metal ion centres through a dense thicket of protein atoms, with the aqueous solvent being almost completely excluded. The interplay of experiment and theory has been particularly fruitful in the area of ultrafast chemistry as highlighted in the review by Bagchi (**page 129**).

P. B.

CURRENT SCIENCE

Volume 69 Number 2 25 July 1995

CORRESPONDENCE

Patenting biotechnologies

The National Institutes of Health (NIH), Bethesda, is seldom away from controversies where patents are involved. Perhaps, as a major sponsor of biological research, it stands to gain more than any other corporate body in the USA. Bernadine Healey, the former Director of NIH, first pushed for patenting of random human gene stretches. These were dropped in a hurry when the scientific community protested. Even Medical Research Council of UK after beginning like NIH soon backtracked. Last month, however, NIH volunteered back into the hotseat. And this time, it is for patenting the basic technique of *ex vivo* gene therapy! The application made in 1989 had initially sought protection for all techniques of gene therapy, both *in vivo* and *ex vivo*. The claims were watered down by the US patent office and now the patent covers only those *ex vivo* manipulations in which malfunctioning human cells are genetically altered to produce therapeutic levels of protein outside the body and then replaced. Fortunately, this patent does not cover 'alternative techniques' and is restricted to the United States. Even then, the cover is big enough to alarm researchers. Normally, this would not hamper research. Yet anyone wanting to commercialize any form of gene therapy may have to get a sublicence to do so. With many players in the game, it does not appear likely that the patent will go unchallenged. There has been of late some interest in gene therapy in India, and the US experience is a pointer for Indian professionals engaged in biotechnology.

What we Indians should especially take note of is the receptivity of the US patent office to changing technological and commercial scenario. The move by the US PTO underlines its realization that gene therapy is likely to succeed as a viable therapeutic technology in the coming years. The departure from its previous stance of insisting on extensive proof of the utility of the invention before granting a patent only strengthens this observation.

The Indian patent system, in comparison, has twisted itself into knots making compromises. Notwithstanding the failure of the bill in the Rajya Sabha and its dubious fate, the Ordinance does not have any teeth for biotechnology inventions. The current stance of the Patent Office to turn down anything genetic or involving DNA ('Agracetus syndrome') will do more harm than good to Indian science. There are several patentable inventions in the country's many laboratories that are not being patented abroad because of financial constraints, nor can they be patented in India because of the stance of the Indian patent office.

What laws apply to biotechnology patents in India? The Indian Patents Act of 1970 states that novelty, utility and a manner of manufacturing are the prerequisites to patenting. When a tangible substance or product results, the process can be patented [see 2(1)(J)]. Unfortunately, the Patent Office does not recognize methods of diagnosis, recombinant organisms, etc., as a substance or a tangible end-product. Read together with section 3(i), which states that 'any process for medicinal, surgical, curative, prophylactic or other treatment of human beings or any process for a similar treatment of animals or plants to render them free of disease or to increase their economic value or that of their products', the scope of obtaining a patent on biotechnological inventions in India is highly restricted. A crucial issue we have not addressed so far is defining microbial organisms in the legal context. Patents on recombinant organisms, elsewhere, are adjudged by the amount of human ingenuity involved and this calls for a legally defined reference point. Unlike other advanced countries, we have not made any distinctions between microbes and higher forms of life. Consequently, microbes are treated at par with animals and plants, and any method of modifying them contravenes the directives implicit in section 3(i) and therefore is not patentable. DNA probes, cloned and genetically engineered organisms are thus not patentable in India. Any biotechnological process done *in vivo* (even in a microorganism) does not lead to a new 'substance' while also enhancing its economic value, thereby rendering the invention nonpatentable under the 1970 Act! For food and drugs, effective protection will still be 5 years, a disincentive when one realizes that drug development may take at least as long, if not more.

There is no concordance amongst the various patent offices of the country in their interpretation of inventions pertaining to biotechnology. There are instances where applications disallowed by one office have been admitted in another. To compound to that, it is understood that an internal communique from the Controller General forbidding patent protection for certain biotechnological inventions, assigning no grounds for such a stand, has befuddled patent attorneys. This obviously has no legal sanction under the Act, yet it is another potential obstacle for the patent seeker, who then has to resort to legal recourse.

Filing applications under the World Trade Organization (WTO) has been seen as an alternative. Product patent can no doubt be sought under the WTO, yet biotechnological issues are, in effect, sidestepped. The ordinance does away with Section 39*, which made it mandatory for an Indian inventor to first seek an Indian patent before applying anywhere in the world! Under the new regime, an identical application must first be accepted in a convention country after 1 January 1995 against which the inventor can seek exclusive marketing rights in India. If the invention was made in India, and a process patent for producing that particular 'substance' had been granted, with the approval of the Controller, the inventor can enjoy exclusive marketing rights for five years or till a patent is granted or the application rejected. In all other cases, the Controller would not be calling for a patent examiner's report till 31 December 2004. Thus, even filing under the WTO in India means a long cold storage, a not too satisfactory proposition when one knows the short life-span of biotechnology products. Though product patents can be obtained for drugs, medicines and agrochemicals, the amendment does not affect the existing restrictions on the grant of patent protection in respect of *substances produced by chemical reaction*, alloys, optical glass, semiconductors and intermetallic compounds.

Indian inventors seeking patents abroad are very often stymied by the substantive examination of the US PTO. In determining prior art, the US PTO scans the available literature, which is apparently disjointed with no obvious interconnections, and then mosaics them, to state that from the emerging cited literature the idea is a logical extension and, therefore, obvious. The other stumbling block is the protection of patents and their 'alternatives'. Any substance substantively made by a similar process having substantively similar properties and giving substantively similar results is deemed as an 'alternative'. The cost of clarifying and/or modifying claims in response to the examiner's report is also too costly to discourage further action.

There are several Indian laboratories doing good work with genetic probes. Since patenting gene sequences is not allowed and any diagnostic process is deemed unpatentable, many Indian researchers are losing out in the international race.

The new use of a known drug has again become patentable in the US. With the expertise available in the West to isolate, characterize and synthesize plant alkaloids and other natural products with therapeutic properties faster, there is a potential threat that soon our wisdom of traditional medicine will be rendered impotent.

So what corrective measures do we take? Primarily, the ambiguity of the new Act must be resolved. Scientists and legislators should come together to propose distinctions between microbial entities and higher life forms on a legal perspective. The cost of patenting abroad is discouragingly exorbitant. Having gone for patent harmonization, would it not be logical to offer Indian scientists a levelled playing field to commercialize their inventions? Can we not make it possible for them to patent their inventions in India speedily so as to obtain a priority date to beat their competitors? It is not enough that we impel our scientists to patent. It remains with our planners to make the patenting environment conducive. We tend to make our laws not by anticipating trends but by following footsteps, even when they do not lead in our direction. It is time that our law makers took cognizance of Indian research capabilities in the biotechnology, pharmaceuticals and agrochemical sectors and legislate to favour them in the troubled years of international competition. We have been for far too long, as a nation, walking forward with our eyes over our shoulders!

SURESH CHANDRAN

National Institute of Immunology,
Aruna Asaf Ali Marg,
New Delhi 110 067, India

*Section 39 is back, with the lapse of the Presidential Ordinance.

Why not platinum?

This note is intended to draw the attention of geoscientists to the need of looking at the platinum content of continental sulphide deposits.

The discovery of attractive quantities of gold and silver in oceanic sulphide deposits prompted a renewed interest in the continental sulphide deposits. Several important and interesting papers on this aspect were presented during the Annual General Meeting of the Geological Society of India in December 1994. Surprisingly, none dealt with the platinum content. In fact, one of the first publications on ocean floor massive sulphide deposits by Hekinian *et al.*[1] did report on the platinum content of the East Pacific Rise (21°N) deposits. Electron microprobe analysis revealed that platinum is present as a dispersed constituent in most sulphide phases that contain gold and silver. Its concentration ranges between 0.1 and 1.4% by weight. Further details of platinum in different sulphide phases can be found in Hekinian *et al.*[1].

Studies of oceanic sulphide deposits have been important in providing insights into the genesis of continental sulphide deposits which themselves formed in an oceanic setting once upon a time. So, when oceanic sulphides contain attractive quantities of platinum, why not look at the platinum content of continental sulphides, in addition to gold and silver?

If good amounts of platinum are indeed present in land deposits, it can change the scenario tremendously. One may also wish to have a fresh look at the platinum content of tailings from different sulphide deposits in the country.

1. Hekinian, R., Fevrier, M., Bischoff, J. L., Picot, P. and Shanks, W. C., *Science*, 1980, **207**, 1433–1444.

R. SHANKAR

Department of Marine Geology,
Mangalore University,
Mangalagangotri 574 199, India

The Academy document: A perceptive appraisal of science education in India

This is in response to the article entitled 'University education in Science' published in *Current Science* (10 February 1995), which emphasizes the need of an effective national policy on science education and research. The Academy paper is timely specially, because, of late, there has been considerable debate on the proposal made by a nonresident scientist, S. Mahajan, to set up a National Science University (NSU) in Delhi. The proposals included in the Academy document, unlike those in NSU, appear to be realistic, goal-oriented and remedial as regards attempts towards improving the overall status of science education and research in India.

The second part of the report that documents Academy initiatives is laudable and unique. The recommendations made under subheads such as 'efforts by the Academy through the fellowship', 'interaction with student and teacher communities', 'link with teaching departments and institutions' and 'contact with national bodies, agencies, etc.', if implemented properly would indeed elevate the level of education in science. However, I am of the view that the panel while making recommendations has avoided important issues that may have a bearing on the successful implementation of the Academy proposals.

The panel which has prepared the draft is surprisingly silent on the issue of what the medium of instruction in science education should be. In some part of our country science teaching at undergraduate (UG) as well as at postgraduate (PG) level is conducted in the vernacular. For example, at PG level in biology, when a teacher starts talking about the pituitary gland in a classroom, one should not be surprised if he finds a couple of blank faces staring innocently expecting to know the meaning of the pituitary gland; or in a physiology class a student might gather enough courage to ask, 'Sir, can you explain what do you mean by kidney?'. I am sure none of the Academy panelists has ever faced such a situation in a PG lecture room. Only a single line '... production of science programmes using the electronic media, possibly also in regional languages ...' has been written in the form of a recommendation on this issue under the heading 'Academy initiatives'. The students usually do not have the ability to comprehend science literature, which is mostly published in English and one can only imagine the situation when one of these students graduates and becomes a teacher. We must think of the ways and means to break out of this cycle, which is a major cause of deteriorating science education. It is, therefore, better if we do not talk about equity and excellence prior to formulating a concensus national policy on the medium of instruction for science education. The English language has been widely accepted all over the world as the language of science. Why should we not openly and boldly recommend English as the medium of instruction starting at least from the UG level throughout the country. This will definitely help elevate the level of science education and research in our country.

The existing age-old examination system is absolutely poisonous for the growth of science. The way the theory and practical examinations are conducted is extremely dangerous. In fact, science students have become habituated to go into orbit effortlessly. Objectivity in test procedures should be brought in and somehow the feeling of 'great' about science subjects should be infixed into the minds of science teachers as well as students. Furthermore, it is well accepted that the absence of accountability makes the examiners extremly casual and callous while examining the answer books. A transparent examination system may be helpful in alleviating the distrust of the student community in the system. On demand, photocopies of the answer books may be provided to the students. This procedure would definitely induct accountability among the examiners and would make the examination system much more credible.

In specific areas like biology special problems exist. Students of biology get their M Sc/Ph D degrees mainly in subjects like botany, zoology, biochemistry, bioscience/life sciences. There are many more, such as microbiology, biotechnology, physiology, genetics, environmental biology, etc. However, in some states step-motherly treatment is accorded to the students with Master's degrees in Biosciences, Biochemistry, etc., by the concerned Public Service Commissions (PSCs) at the time of recruitment of lecturers for hundreds of colleges, where only botany and zoology are taught. The University Grants Commission (UGC) favours introduction of new courses but never cares to protect the products. This cold war between the UGC and the state PSCs has been continuously ruining future careers of hundreds of students every year. This type of discrimination may dampen the spirit of those science students who opt for new branches of biology, thereby rendering modern aspects of the subject less attractive.

The Academy while constituting the panel should have co-opted some members from state universities, remote universities, and colleges. I feel that the Academy panel itself is totally skewed and has 7 out of 11 members from Bangalore alone. The technique of random sampling for the nomination of panel members may have been more fruitful in pinpointing the real problems faced by thousands of science teachers spread over the length and breadth of this vast country.

ATANU KUMAR PATI

School of Life Sciences,
Pt. Ravishankar Shukla University,
Raipur 492 010, India

The Academy document: A teacher's view

It is heartening to learn that the Indian Academy of Sciences has prepared a Paper on 'University Education in Science'. The paper outlines many steps which the Academy plans to take to improve the quality of science education in India.

The Academy is, by its very nature, an elitist organization of those involved in research in various science areas in the country. Its Fellows are held in high esteem. But I am not aware if any one doing frontier work in the field of science education, *per se*, has been elected a Fellow of the Academy. In that context I feel doubtful if the sudden interest shown by the Academy in the field of education can be fruitful, unless they plan to take the universities and science teachers into confidence at the level of equals.

Before independence, science showed a healthy growth both in teaching and research. This was primarily because the universities were the centres of attention in both. With the rapid growth of many specialized agencies for research in science, with large funding and relatively little external monitoring, the universities came to be neglected, and in successive stages they were emasculated both in manpower and funding. That is what brought the fall in standards in science education. If the Academy is unable to see this background and suggests a reversal of the process, no long-range improvement can be brought about.

As matters stand today, every time an idea to improve something comes up, we tend to think in terms of creating a new island in the ocean of education. The interuniversity centres form one big example. The thought of advanced centres and elite institutions stems from this line of thinking. The Planning Commission report which the Academy has extensively quoted, and to which the Academy Paper appears to subscribe by and large, is on the same line. The reason is simple: the Planning Commission could not identify 'good enough' teachers from the universities (just as the NCERT, for example, fails to find such people from the universities) for consultation, and the result is the report which the Academy Paper quotes in depth. Even as such proposals were afloat, the best of universities gave up undergraduate teaching (and the worst of colleges switched from 'general' to 'honours' courses at B Sc level). In effect, the centres of excellence tended to treat undergraduate teaching as something below dignity.

Limiting myself to physics, I find that the Academy Paper fails to make even a mention of the existence of the Indian Physical Society (IPS), Indian Physics Association (IPA), and Indian Association of Physics Teachers (IAPT) as organizations with interest in the advancement of physics education. IAPT, in particular, has been publishing a monthly journal dedicated to physics education now for over 10 years, and yet, when the Academy Paper proposes publishing a new journal it does not even mention that one such journal exists, and four journals initiated and funded by UGC for this very purpose are also in existence.

The UGC itself has a Panel for each subject, like the Physics Panel presently under J. V. Narlikar. The Panels are supposed to do exactly what the Academy is now planning to step in for. I wonder if the Academy ever thought of consulting or contacting any of the UGC Panels. As one example, for UG physics laboratory, the Physics Panel conducted a Workshop at Mangalore in 1994, which does not find any mention in the Academy Paper, although efforts by individual institutions are listed there. IAPT in 1992 had moved that a Workshop to sort out a Model UG Lab in Physics be held to identify the most educative open-ended experiments, and it gave an initial list (to begin with) after a three-day seminar. The present Physics Panel of UGC has moved to set this into action, and then have it multiplied at 10 places in India for year-round training/orientation of teachers. The Academy Paper does not show that its Panel members are aware of such a move. What they suggest shows utter ignorance of how educational laboratory items are developed. Having been involved with the work of Saraf for over three decades, I know it needs sustained and exclusive attention over many years to get somewhere. Just sitting across a writing table and producing a list of 100 experiments amounts to nothing more than duplicating what the Boards of Studies at some 150 universities have been doing for the last 50 years; it could be worse if the persons involved are leading researchers and not teachers with a special eye.

There is also reference to equity and excellence, whether it is to be *and* or *versus*. This is the reaction of those in the upper atmosphere, who are unable to see the ground situation in science education. Firstly, I see no point in dabbling with a subject which is not my own, viz. social science and democracy. These have their own pressures and consequences. Our decision to improve the quality of science education must in no way be linked with what happens in the social field. Secondly, and more importantly, we must see that science education has three clear objectives: (a) preparing candidates to take up higher education and research in a chosen science field, (b) enabling them to take up various applied areas where science is to be of direct use, and (c) creating general enlightenment and enrichment of mind to be a better citizen of tomorrow. As a recent survey by IAPT has shown, not more than 10% of those who enter B Sc have objective (a). This is nothing to be surprised about. Most of those seeking professional courses have been sorted out, and those that remain enter B Sc largely for objective (c): enlightenment and enrichment.

Here comes the major distinction between how this analysis would be seen by Fellows of the Academy and by teachers. The former would have their eyes rivetted to only students of kind (a), while we, as teachers, have as much, if not more, eye on those whose objectives fall under (b) and (c). This brings me to what has happened in the past few years. At one time the UGC Physics Panel was exclusively manned by non-university personnel. The Curriculum Development Committee of the UGC was manned largely by university teachers, but their eyes were exclusively on kind (a) objective, partly because they had to create a syllabus that should look attractive to top physics people, which means research people, not mere teachers, like me. And with the passage of time, university after university went on revising its syllabus to make it look the most attractive in the same sense; worst, the distinction between a general and honours physics course was wiped out.

Typically, out of some 12 papers of the same trend, if 8 were taken it became 'general', and if 12 were taken it became 'honours'. The difference in the approach and spirit of the two, which is linked with the difference of objectives, got lost. The result has been that the general student is left high and dry, as also the teacher.

Even worse happened at the higher secondary level. The NCERT, mistaking the best research workers to be also the best teachers and educationists, invited several top elite people to head committees to write the syllabus and text materials. I do not at all cast any doubt on their honesty in such selection. These eminent people accepted the responsibility, and I again do not have the least doubt that they did mean doing good service to the society in all honesty. They are all, in my opinion, men of great honour and competence in their respective fields. What went wrong was that they did not realize that education is a separate field of specialization, requiring different talents than only knowledge of science. Further, they were not prepared to give up their current research involvements, and they left most of the work for others to do. Yet worse was that once the manuscript was handed over by them, they did not have control on the get-up, drawings, conduct of examinations, etc. This cascading series has done very considerable damage to science education in India. One has only to compare with the effort of an agency like the Physical Sciences Study Committee in the USA to find how very poorly we have come forth in this effort. I mention this in depth because the Academy appears to be entering the same kind of role, in all honesty, of course, but without quite realizing that education is a different field from research.

Let me repeat that I have the highest respect for the Fellows of the Academy for their contributions to research. I also have the confidence that the proposal they have come up with springs from the bottom of their heart. They really mean to help education in science. In that context, I have a few suggestions to make:

(a) Let them not start a new journal in science, but send their articles and suggestions for improvement in science education to the existing journals. Their frequency can be increased; their volume per issue can also be increased. In fact, even their format can be altered. I am sure the Editors of the current journals will be only too happy to have many of the contributors on their Board of Editors.

(b) Let them use the channels of the UGC subject Panels to initiate new ideas and proposals, rather than create a new power centre of their own. Let them also use subjectwise teachers' organizations as means of executing their programmes. For physics, from IAPT, with over 3000 members spread uniformly all over India, and with its sustained work over 10 years, they can feel sure of all cooperation.

(c) Let us discuss in depth why physics is losing all interest and excitement as a subject to be learnt. In particular, the laboratory component needs to be examined in depth. (But it cannot be a casual matter.)

(d) Let them involve leading teachers from the universities (as distinct from leading researchers only) in whatever they want to do for the advancement of science education.

(e) Let the very idea of creating elite centres for elite students be shelved deep, and let us work for the uplift of the largest masses through dedicated teachers' efforts. The processes of examination, selection of candidates as also the institutions, and the selection of teachers, etc., are all fraught with great danger. They may lead to more fragmentation in the society, and particularly the teachers.

I appreciate the initiative taken by the Academy. The purpose of this correspondence is to get the best out of such a powerful initiative.

D. P. KHANDELWAL

B-31, Koregaon Park,
Pune 411 001, India

OPINION

The Indian Academy of Sciences document on science education in universities: A criticism

V. Sitaramam

I was dismayed when I read the action plan of the Academy for science education in Universities. It is necessary to assess what the Academy proposes *vis-à-vis* what its predecessors have done. The document rightly refers to the Kothari Commission, a large tract that bemoans the state of education even as it stipulates what needs to be done. The Academy document, more prosaic, more clinical and consequently less transparent of its own shortcomings, has no insights. If it has, these are suppressed very successfully[1].

There are three players in the educational scene: the teachers, the students and, most importantly, their homes. Each of these has undergone a sea change in the last few decades the world over. The document reflects neither their magnitude nor the implications. Most importantly, the document offers no justification as to why science education has to be given specific preference. It has some marginal statistics, as a matter of numbers. The reasons behind the numbers are anybody's guess. If

students are not taking up science, what are the reasons? If they are taking up science and not doing well, what are the reasons? Instead of going into a factual analysis of what is the situation, the document considers the simple-minded scenario: put in money where the students are good, the results are bound to improve. A number of instances are cited as patent successes such as the Maths Programme and the Biotechnology Programme. The document does not reflect any of the existential concerns that affect these so-called successful programmes.

So the major concern is money. This country has developed two major arguments in defence of its failures: first that there is never enough money and the second that there is too much population. Education and science are no exception. Good conscience for a scientist, for a Head, for a Vice Chancellor is to get money. Many of our mind-sets are based on expensive fallacies, which are more important to understand than what the Academy has to offer. Let me briefly outline these.

Fallacy of money

Give better wages to teachers and teaching improves. It did not. Easy research money also will not improve research. In an excellent evaluation of an Indian university in an anthropological study, Don Kurtz wrote of how majority of the rural and a great deal of urban teachers are busy with parallel occupations so engrossing that their teaching is a mere avocation. The shocking thing about the university teacher is not that he attends to other jobs. A good many do not. Majority spend their time vacuously without any yardstick of achievement except routine teaching. Someone should do elapsed time analysis of a university teacher (this is more difficult to assess in National laboratories, where timings are adhered to but there is no rule that one should work!). This easy style of living is reflected in the low level of maintenance of buildings and rooms.

It is a sad mistake to imagine that one is complaining about all these. Actually it is a very comfortable situation. There is total lack of accountability. Good manners and departmental *esprit de corps* see to it that no one is challenged. Administration is usually very accommodating. Each faculty member has enough reasons of committees of various kinds, politics, harassment and what not to prevent him from teaching or doing research[2].

Fallacy of perfectionism

This is the major bug bear of Indian educational system. I think that it is time we realize that a great deal of standard teaching requires plain donkey work. If it takes donkey work to teach at the lower levels *en masse*, then either we should be prepared to do donkey work for majority of the time or we should not be ashamed to appoint donkeys to do that. This is what work ethic is all about. Kothari wrote about the requirement for a university teacher. He wrote, in a sense correctly, that a one-hour lecture should require 3–4 hours worth of library and assignment work by the student. That is not easy! If I have to give a *thinking* lecture, original in content (as opposed to an information-laden lecture), based on what I know, without much preparation, I have at most about 5–8 lectures to give per year! Otherwise, I have to resort to existing problems in books or reading assignments to make them work.

If we consider science of today as perfect science, we already lost the battle. We would teach science as a finished polished perfect thing. It loses all the uncertainties, imperfections, greyness of the most exciting part of science, i.e. the developing science. The most damaging aspect of education is that most teachers take the finished aspect of science as gospel truth and consider their job primarily one of giving information. Science cannot be taught without doing it. Alas, it also cannot be done without sharing it. This was why most argued that the universities should be the places for research, a mould we have thoroughly destroyed in this country.

It is not that there are no problems. If we teach science as the grey imperfect thing that it is, we lose most of our audience. A sustained effort at teaching the method of science to M Sc Biotechnology students over 5–6 years convinced me of the futility of it to get it revoked from the syllabus! It is not possible to teach something in isolation when it is not shared by bulk of other teachers.

Fallacy of merit[3]

This country started IITs with the hope that if we take the best and put them where they do better, everything will be alright. What have they done after IITs? If intelligent people need to be doing science, why is it that so few of our brethren in IITs have impressed these growing and glowing minds towards fashions and passions of science and research? Why do many opt for the Citibank? The students who get into IITs by grind do so under a number of compulsions, parental pressure being the most, peer pressure the next. An individual's choice for a subject is dictated not by active interest but by the current mores and the order of merit. Merit does not seek personal involvement. What has recognition of merit led to? The evidence is clear that any recognition leads only to easy life. So does lack of recognition, which blunts ambition. What we fail to see is that by sequestering them into a privileged community, we have raised their expectations without a provision for integration.

Fallacy of preaching

Most of us think that since we are doing science, what we are doing interests our students. There is also this law of diminishing returns in science. If you have done science, if it is the best thing for you, so be it. If an area has rewarded you, it need not interest the discerning youngster. The student wishes to examine his returns. Secondly, he wishes to know if you work (≡ succeed).

Since most do not work by themselves in the absence of students and assistants, and many do not work regardless of students or assistants, the question of science in this country is limited to what can be done by way of Ph D programmes, which are primarily remote-control operations. We are talking about a very inexperienced work force, whose life is limited to 4–5 years, of which 70–80% is lost at the formative period and hardly 10–20% of the professional life goes into research. If we consider countries with much lesser number of people, they have achieved much more. The difference is not in overall numbers but in working numbers. We have the least of the latter. These cannot enthuse the newer entries towards anything other than a retired life.

Fallacy of rewards

Any reward system that is not heavily backed by peer pressure is a disincen-

tive rather than incentive. Scientists would be no more than dictators of tin-pot republics displaying rows of medals that were not won in any battle and were largely self-endowed. Even then, the real danger is not in the awards but in the deliberate creation of pecking orders. Lack of authenticity, the main plague of Indian science, leads to many alternate means to develop science and education, all of which are doomed to fail and they have. The best of the Indian institutions are imitative. In the absence of originality, and even more, authenticity, where does excellence come from?

There is something about the past and present in the Indian scene of science and education that is worth pondering over. The key word is authenticity. I refer first to the Kothari Commission and then to the psychology of the modern student.

Fallacy of degradation

One discovers corruption and degradation as part of the middle-age crisis. Is our science deteriorating? *Or is the middle-age science deteriorating*[4]? I would strongly recommend a reading of Cicero to anyone who talks of contemporariness of corruption and degradation.

When we are younger, we do see that things are not going too well. Being younger and optimistic, we let things be and get busy with whatever we can do. We soon run out of time. The crisis of middle age is when we are running out of new things to do and running even more out of the time available to us to find new things to do. A clear case of burnt out options! When we find that time is finally getting the better of us, the illusions fast disappear and we despair of the things that would never be. How well does the government recognize science? How about a new NRI university? We forget that we represent a constituency and a trade union and we forget that we have vested interests and we also conveniently forget that we have larger responsibilities. Relevant work in this country is shunned since it involves commitment. Why do we blame our politicians for monopolizing the public interest? Do we share any? It is possible to live with all these imperfections among ourselves. But the instant we arrogate ourselves to the keepers of conscience, we certainly have gone into a severe level of delusion. Too many of us sit on too many committees too often as yesmen. We not only achieve little but also allow things to slide far. Kothari was deeply worried about this docileness, this lack of comprehension of larger scheme of things, this apathy and cynicism and the lack of ability to challenge the purpose.

Kothari's forgotten plea

The conceptualization of a modern Indian university in the Western mould was literally a thought in the wilderness. It was stated that universities control colleges that are not even in the same province. It was contrasted with Oxford and Cambridge and the British thought that one day these problems would disappear when colleges do not depend on remote control by distant universities.

The UGC came into being and a lot of universities followed. All of it was inevitable and did not worry Kothari. One thing worried him most. It was the independence of thought. He was deeply concerned about the fearlessness of the Indian teacher. In fact, this is among the few personalized and important topics that he devoted most attention to.

Obviously, what Kothari was worried most about was the lack of individuality in the teaching community in this country, because that is what he wrote about. He eloquently asked for it since this is what we did not, do not and will not have. We actually confess this gently all the time. When we say that we have failed to develop adequate peer review structure in the country, do we not mean that we have become mercenaries, with no love except for our own gains? We at least admit our lack of honesty of purpose.

I take it that sycophancy was what worried Kothari most. The British loved it. The colonial India had its ethical roots in Rudyard Kipling and sycophancy[5]. Independent thought was possible. Fearlessness was most definitely not[6].

Kothari's errors

The Kothari report wielded considerable influence. We had M Phil degrees as a further stepping stone between Masters and Ph D. There was this pervading mind-set that, in order to do something new and original, it requires proper preparation and that this preparation can be imparted by the wise and holy. It is now common knowledge that even a Masters degree is a waste of time if the B Sc is okay. The way to learn to swim is to start swimming and not prepare lecture notes on swimming. So also in research. What do we find about M Phils? These are either clocking time to go abroad or get married or indecisively waiting for better opportunities like MBA or IAS. The right approach to research is to get going and not to prepare for it.

The second error of Kothari was in that he romantically believed that production of some high-powered, expensive structures for education for the meritorious would enhance the capabilities and the nation would benefit. Yes and no. This class of problems has at best a percolative effect. This contributed more to the baseline than to the top performers. Nation building requires a lot of chores that may not have much market value. For instance, who would go to regional colleges to teach? When the USA stopped the preferential import of engineers, a lot of them stayed back and now we have some Ph D's teaching in colleges, which was never the case. Other than a marginal percolation, have these edifices contributed significantly (considering the money poured in) to the growth[7]? The error is not in asking for merit-oriented establishments of high cost and fidelity in performance. The error is in assuming that such structures would be possible in a time-lasting manner and that the results would be independent of the social circumstances and market forces. The kids who study there come from families and families have aspirations. The question is, do they match?

The major problem lies in the romantic notion of merit. Merit yields results (for the individual) and results lead to power. The shortcut is that let merit be assumed and not won. We have only tenured positions. Who cares whether a qualified person does *not* work or an unqualified person does *not* work. The end result is the same. Universities have offered us exactly this. This heady brew of nonperformance. There is no academic audit in any university. What is in question is not merit but the attitude to the job at hand. Merit does not ensure that and, therefore, it is not a sufficient condition for performance. We have not evolved a system that ensures merit as well as attitude.

The market force logic

When we were young (I grew up in a University Campus), everyone told us that it would be good to take up a professional course since jobs are easier to get and assured income is what life is all about. While that may well be, the academics were also seen to be doing things on their own and appeared to be equally comfortable. While emphasizing the material things, there was a simultaneous de-emphasizing of their primacy in life, we have not heard in schools and colleges anything other than that education is everything and everything else is of less worth. A mind-set that persisted by frequent repetition. The market force was always there[8]. Only, it was not put on a pedestal like now. The best always exploited the best of the market since they are both savvy and have highly developed manipulative skills (as any teacher would know). By de-emphasizing the market forces, the second best, though with higher motivation, got into education and related areas, including government jobs. Now the industry, which is still considered second-rate in the country, is picking up these people vigorously. So enrolment in science will plummet further since new markets are opening up for them. How many are taking language and philosophy? Do we need any lesser brains in these areas, however few may be required?

What do we do today? In biotechnology, we are supposed to encourage students to go to industry. Since we speak of industry, the students ask whether they can have the cake and eat it too. Can they finish the Ph D while working for an industry and so on. Amazingly, few students think beyond the self-reference that they owe something to the industry that is going to pay for their upkeep. The Indian student is brought up in such a sheltered manner that no one told him that there are no free lunches in the free market economy. Little emperors!

Whatever we do for Indian science, elitist attitude will destroy performance. It takes away opportunities from those who need it most and fritter it away on those who have the least commitment, statistically speaking. Science and knowledge are based on performance. Merit is based on promise. These are not the same thing. Power is based on actualization by simply being there i.e. where it counts. Merit and performance are routes to power. Neither individually nor both together are sufficient conditions for acquiring power. Amongst the proxy routes is what we call nepotism, a word we reserve for the individual. When a group is involved, we call it social action and equal opportunities employment. There is no point in getting our hackles raised. All these explicity recognize that being there is what matters in society and, therefore, in the absence of critical performance criteria, everybody will insist on equal opportunity to get there.

Teachers, the romantic diehards

Pune University has an interesting experience, since it housed till recently zoology and biotechnology in the same department, the former by local merit and the latter by national selection. As an insider, I tried to evaluate what motivates the teachers the most. The biotechnology students represent, on an average, a more cosmopolitan cross-section and perform better. However, one needs to listen to the teachers about their attitudes to these students. The university teachers, unlike the guest lecturers from national laboratories, react differently. First of all they complain often that these students are less sincere and more arrogant. When the same complaint comes too often from even the better teachers, one worries. What I saw was an amazingly simple matter in academic psychology. The faculty claim that while a lot of teaching is drudgery, the tedium is made up by the few students who perform very well. The usual university students tend to be of uniformly lower level, barring these occasional performers, and their motivation also tends to be uniform. In this low-average class of performance, the occasional good student could bring tears to the eyes. Of course, it is not as simple as that. Actually many of them are exceptionally nice kids, often very sweet and most polite. They are quite diffident and yet when handled gently can be motivated to a higher level than one would consider it possible. On the other hand, the biotech students are, on an average, better in their intelligence, with less of an excuse for goofing, which makes the teachers mad. More than 50% of them have no interest in what they do once the novelty of Pune life wears off. Thus, we see better performers with a higher level of variance in their motivation and performance among the biotechs. It is definitely more painful to see a better student being a nonperformer than a lesser student. Since such a possibility arises more in biotechnology than in zoology, the attitudes of the teachers too have shown variable level of weathering towards these students. Better students breed cynicism faster in teachers than lesser students. The culprit is the variance in attitudes at constant capability! This was confirmed by some IIT faculty as well in personal discussions.

We have occasional departments in the universities with considerably better cross-section of teachers, at least on an average. The student inputs are no better. However, the teachers generally do not innovate in teaching simply because this is a bother. That is what worried Kothari. This oppressive (conventional) wisdom that aborts action even before it is contemplated!

The psychology of nonperformers

In the final analysis we need to contend with nonperformers. Lack of intelligence does not correlate well with nonperformers any more than intelligence assures performance. The key factor is the attitude. What drives people to nonperformance? Are they made or are they bred?

We cry a lot about changing values and loss of traditional modes of living for all the wrong reasons. In the larger joint families, the average attention on a single child is negligible. There are so many. Largely the children look after themselves with marginal contact with adults. What happens now? With the increase in nuclear families, the child (we two and ours two!) now has two pairs of grandparents and one set of parents who dote on the single or the well-spaced two children. The parents, both working, come home with a heavy guilt feeling that they are neglecting the children. This view is heavily reinforced by the ageing grandparents, whose justification for hanging around is that they are needed to fortify (or disable?) the minds of the younger with their notions (when we were younger, we rarely spoke to the elders and, fortunately, they spoke to us even more rarely) since the parents are neglecting the children. The guilt complexes lead to overindul-

gence of the children. Thus, little emperors are born. Once born and bred, they stay! China is reeling under the impact of these little emperors. I have interviewed a large number of students now. I have no doubt that we have our own growing numbers of little emperors.

We have been devoid of performance in this country for a long time. There must be deep-seated reasons here for this. Individual excuses or else a social reason like population or caste cannot be made responsible for it. *We need to consider it a matter of highest priority as to what makes our youngsters untick. We need to determine the behavioural strategies we adopt with great care as parents, teachers, grandparents, or what have you, that rob the children of their initiative*. There are many things that we do not know about our youngsters. Most of all we do not know how much of insecurity is critical and undamaging to the youngsters to ensure their performance. We need to consider even their social mores, driving force for smoking to harder habit like drugs, AIDS etc. Unfortunately, there are few meaningful studies on a cross-section of people that represent tomorrow.

What we require is a professional psychological survey aimed at the performers and the nonperformers. In spite of the fact that I believe that I have seen a reasonably large sample of semi-performers (as in biotechnology as well as in IITs and so on), I feel it is very prudent to have some assessment of what is going on. We need thumb nail sketches of what the parents think. We need to know what made the students opt for what they did. Above all, we have to examine why the Indian science student has delayed social puberty. I have seen several students react to their first time out-of-home atmosphere as if there is no tomorrow for catching up with all the lost time. We have social dimensions and problems of a large kind that determine the behaviour of the youngsters.

This is where the notion of perfectionism hurts. Science is perhaps the only avenue in academic life in which a mediocre person can do well. Arts and humanities are brutal. To do well you have to be very good indeed. Instead of emphasizing performance that we should, we emphasize promise. There is no other way to understand the Academy document. Performance requires motivation and motivation is in the realm of psychology, an aspect of human beings that the Academy was deafeningly silent about.

The approach of the Academy is not without blinkers. First we have those who take up science because they have not been pushed enough to get into competitive exams. Then we worry about their future since we find the options diminishing. A lot of ideas get generated: NRI universities, new initiatives for science and so on. In all this we ensure the basic attitude that spending money is good enough.

The action programme by the Academy: How justified is the fund allocation?

The Academy followed up the recommendations of the Planning Commission and suggested spending some 25 crores on the gifted, 60 crores on the less gifted and another 30 crores on the rest. At the face value, it looks reasonable. Have another look and you would be horrified (Figure 1). The numbers look alright grossly. The Academy is suggesting spending some Rs. 15,000 per capita on some 700 students of the tier I per year, some Rs. 1000 per capita on some 25,000 students of the tier II and some Rs. 100 per capita on some 1,23,000 students – the rest. These being 1989 figures, an upward revision of some 5–8-fold would be in order. I have no doubt that there is no ill will and the recommendations were made with the best of intentions. But then, good intentions are what they are. Where do they lead us? On the one hand, we are talking about 60 crores (tier II) of the same level of incomplete education, while we are offering 30 crores of placebo *as a political gimmick* to concentrate 25 crores on the (?) desirables. We are talking about, with a straight face, creating a yet another group of not-so-little emperors who will also become self-seeking snobs. They would also be alienated and subservient to all that is Western, with a goal no more than to dwell on a postdoctoral fellowship abroad. For a few more dollars! Yes, Kothari had a justification then. The performance outcome of higher citadels of learning towards science education, or even education in general, was not understood then. The age of innocence may be forgiven. Yet we are talking about reinforcing majority of these 'failures' (the Academy was also clear

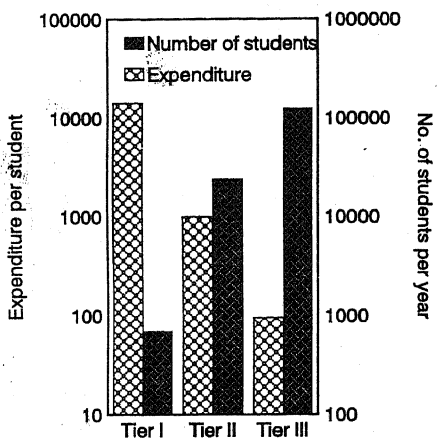


Figure 1.

that these were rejects of an earlier selection process) with a system of education so far removed from their own brethren and for which they have done mighty little to earn.

When they graduate, their expectations will be very high. We have been fighting a losing battle with the biotech students and their attitudes, of levity and lack of sense of proportion. We created an artificial market. It was not bad, insofar as the pressure was kept high on them and they were made to work, such that even routine practicals ran past midnight. When relaxation has set in due to inadequate induction of new faculty, and the change in attitudes with time, we find that even their CSIR–UGC scores have begun to dip. What the Department of Biotechnology has done is losing an excellent initiative, and not help setting the norms. Universities have their own momentum for the inevitable downward slide and oblivion for any programme.

What should the Academy do?

The intelligentsia in this country can achieve more for this country, if they display greater humility. If not, we have these recurring and persistent images of Don Quixotes running at wind mills.

I think that the Academy should continue to promote debate and help provoke people. It certainly provoked me into writing this, possibly much against better judgement. It should have definite and tangible goals that help ideas and proposals for school education, videos, etc. It should also learn from the

existing experience in media management. It is time we analyse and tell the UGC how poor many programmes on the countrywide classroom are. The Academy should primarily help motivate any and all programmes that are based on individual initiative.

Above all, the Academy should stop worrying about excellent science and start worrying about existing science and the existing practitioners. Existing problems are there only because the available solutions do not work. A recurrent problem often requires a drastic change in the mind-set. The Academy also should stop worrying about NSUs. We have built many edifices. There will be more. If we are more worried only about losing the piece of cake, we will have no moral justification for the stand we take. Is all the UGC money going to teaching departments and to equip them better? Have major chunks gone elsewhere?

Is there a definitive, prospective course of action urgently required? Yes, I believe that it is time we get the educational system examined by professionals, including sociologists and psychologists. What we need most is painstaking definition of standards in all that we do and the minimum that we can tolerate. *It is the minimum and only the minimum that we need to define.* The government will not help here since it caters to a different constituency. Only we can help ourselves. Above all, it should also help develop standards for rating performances at all levels. Armed with appropriate surveys and relevant analyses, such an effort may bring out some clarity on where to focus our efforts for maximum results. Another area is to begin a serious exercise in cost-benefit analysis and a realistic appraisal of minimum funding for all universities for given courses. The newly started National Assessment and Accreditation Council has aspirations and plans in this direction, which would require considerable help from the academic community.

Centralization of education was inherent to the Academy's idea of picking a few as the best students in science. It has its evils and is even nontangible due to the displacements involved. A larger goal may be to take higher education out of concurrent list, if we have to survive the competition following liberalization, since we simply have to do better. In any event, the solutions run deeper than realized in the Academy document. It is time we shed the proven unsuccessful paths and exercise the same quality of vision that we envisage for the future students.

Notes

1. Peter Medawar in his *Advice to a Young Scientist* remarked that there have been two preceding instances wherein an advice to the younger was attempted. One is Polonius' advice to his son in *Hamlet*; that gives the basics of noninvolvement . . . 'neither a borrower nor a lender be'. A later effort was Lord Chesterfield's advice to his son, which, aptly summarized, taught his son the behaviour of a courtier and the morals of a whore. In attempting to advice the younger, it was necessary for Medawar to steer clear of the shortcomings of his illustrious predecessors. So he came up with a book that tells you how to excel in science if you are already born with a silver spoon in your mouth!
2. The Academy wrote of conveying the excitement of science to young minds. Pooh! In a class of 100 students or more, teaching the same topic for over 25 years, what excitement does one convey and to whom? Even if one starts well, how long will he sustain?
3. 'I am profoundly suspicious of the "A" man. He can say back what you want to hear in an examination, and' . . . 'you must give him his "A" if he says it back; but the ability not to say the willingness to give you back what is expected of him argues a certain shallowness and superficiality. Your "B" man may be muddle-headed but muddle-headedness is a condition precedent to independent thought . . . ' A.N. Whitehead, *The Dialogues*.
4. Or is that since Fellows of Academy come into being in their middle age, there is a coincidence in the way things are perceived to be wrong?
5. Ramanujan was the ultimate example of colonial legacy. The moral of Ramanujan is not that Indian genius exists. The moral is that the only salvation for the Indian scientist is International science, where he will have a better and fairer chance than in the Indian scene.
6. This is a tradition that is not lost to this date. Rotation of headship (of a department) does not matter. Having spent nearly a decade in a university, I realize that it is merely of rotation of one to whom obsequiousness must be tendered to. A well-practised ploy to a countryman to whom polytheism is deeply enshrined in the heart.
7. It is time someone does a cost–benefit analysis of educational structures the same way as we have done for dams. Taken the overall scenario, the cost–benefit analysis showed that development is not without cost. The cost–benefit analysis should take into account the cost of diverting money and personnel from the educational system and the effects on the dilution that ensued.
8. 'I do not believe that the modern parent of any class is more mercenary than his predecessors. When classics was the road to advancement, classics was the popular subject for study. Opportunity has now shifted its location, and classics is in danger. Was it not Aristotle who said that a good income was a desirable adjunct to an intellectual life? I wonder how Aristotle as a parent would have struck a headmaster of our great public schools. From my slight knowledge of Aristotle, I suspect that there would have been an argument, and that Aristotle would have got the best of it' A. N. Whitehead, in *The Aims of Education*.

V. Sitaramam is in the Department of Biotechnology, University of Poona, Pune 411 007, India.

SCIENTIFIC CORRESPONDENCE

Development of cultured tomato anther to a fruit-like structure accompanied by *in vitro* ripening

With a view to regenerate haploid plants of tomato, anthers carrying microspores at uninucleate stage (buds of 8–10 mm) from glasshouse-grown plants of cv. Arka Saurabh were cultured *in vitro*. The medium comprised Murashige and Skoog[1] inorganic and organic constituents, 3% sucrose and 0.25% gelrite (pH 5.8). Factorial combinations of kinetin

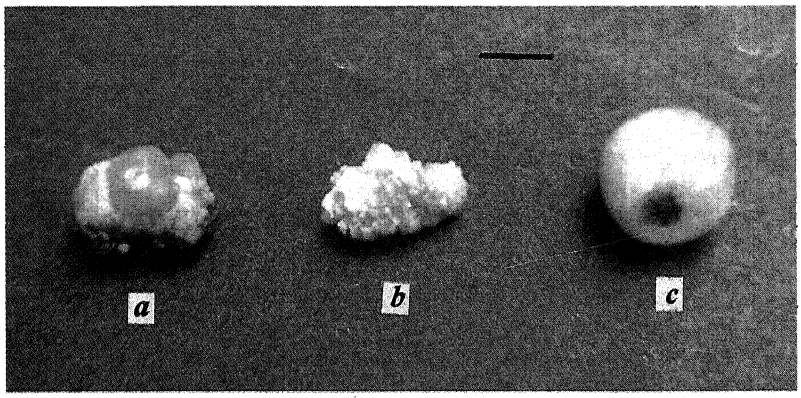


Figure 1. Figure showing pseudofruit at 7 months from culturing of anther (*a*), a half ovary, two months from culturing (*b*) and a normal fruit two weeks from pollination in field (*c*) (Bar = 1 cm).

(0.05, 0.5, 5.0, 10.0 or 20 μM) with equimolar or higher concentrations of 2,4-dichlorophenoxy acetic acid (2,4-D) or indole 3-acetic acid (30 combinations) were used. A total of 32 anthers were cultured per treatment, half of which were incubated under continuous darkness and the other half under 16 h photoperiod (20–30 μmol m^{-2} s^{-1}) at 26 ± 1°C. A limited proportion (30.8%) of anthers showed swelling, of which some exhibited callusing in 3–4 weeks irrespective of the hormonal constitution or incubating environment. Such anthers were transferred to a fresh medium of the same composition at 6 weeks. One swollen anther on medium containing 5 μM kinetin and 10 μM 2,4-D and incubated under 16 h photoperiod swelled up in another 6–8 weeks resembling a small half tomato fruit. Three lobes were there on the upper side and the bottom was flat with reminiscence of calyx-like structure at one end. The pseudofruit on subculturing to fresh medium showed further enlargement (about 20 mm long and 10 mm thick), after which the growth ceased. The ripening process began by 6 months from initial culturing and was a slow process. The top lobe turned fully red while the side ones remained yellow/orange.

In order to verify if the structure resulted from accidental culture of an ovary, a number (200) of split ovaries were cultured on the medium with the same composition. None of these developed to a fruit or a similar structure. Instead, they went into callusing. The pseudofruit was different from normal fruits. While the former was lobed, the normal fruits from field-grown plants were smooth.

Cross-section of the half-ripe pseudofruit showed no placenta or ovule. The inside was full with firm tissue. On the other hand, the normal fruits had ovules attached to placenta embedded in a mucilaginous matrix which was not connected to the outer wall/pericarp (Figure 1). Histologically, the pseudofruit comprised of closely packed parenchymatous cells. The division between lobes was marked with still closely packed cells. The cultured ovaries on the other hand showed a thick mass of tissue in the centre with no distinct cells at the same (× 10) magnification. The outer callus exhibited loosely packed parenchymatous cells. The observations confirmed that the pseudofruit had originated from a cultured anther, with the anther lobes forming the fruit lobes, which is not a typical characteristic of the variety in question. Subsequently, a large number (over 1000) of anthers were cultured under similar conditions, but none showed a similar response, suggesting that this could be a rare phenomenon.

Gresshoff and Doy[2] observed the development of haploid callus cells in tomato to seedless pseudofruits. The present observation is different in that one tomato anther developed into a fruit-like structure. Homeosis, i.e. the expression of one type of organ in the site of another one[3], has been reported in several instances in nature and in isolated cases under *in vitro* conditions. They often appear as mutants[4–7]. Homeotic transformations could also be induced by environmental factors or by interplay of phytohormones. Fisher[8] observed transformation of stamen to ovaries and of ovaries to inflorescence in wheat under short day conditions. Rylski[9] transformed stamen of capsicum flower buds to carpels (with ovary, style and stigma) through treatment with gibberellic acid. Stamen and stigma-like structures, functioning like original units, were induced from receptacle tissue of tobacco *in vitro* in the presence of kinetin[10]. The present observation is different from these in that the *in vitro* cultured anther developed to a seedless pseudofruit just in external appearance and in ripening. No carpels or ovules were present. Such teratological specimens can be useful in interpreting pathways of development. Attempts to reproduce the results were not effective as the specific inducing factor was not exactly understood.

1. Murashige, T. and Skoog, F., *Physiol. Plant.*, 1962, **15**, 473–497.
2. Gresshoff, P. M. and Doy, C. H. *Planta*, 1972, **107**, 161–170.
3. Rutishauser, R., in *Progress in Botany* (eds Behnke, H.-D., Esser, K., Kubitzki, K., Runge, M. and Zeigler, H.), Springer, Berlin, 1989, vol. 51, pp. 54–69.
4. Gottlieb, L. D. and Ford, V. S., in *Developmental Mutants in Higher Plants* (eds Thomas, H. and Grierson, D.), Cambridge University Press, Cambridge, 1987, pp. 1–17.
5. Pruitt, R. E., Chang, C., Pang, P. P.-Y. and Meyerowitz, E. M., in *Genetic Regulation of Development* (ed. Loomis, W.), Alan R. Liss, New York, 1988, pp. 327–338.
6. Valade, J., Maizonnier, D. and Cornu, A., *Can. J. Bot.*, 1987, **65**, 761–764.
7. Bowman, J. L., Smith, D. R. and Meyerowitz, E. M., *Plant Cell*, 1989, **1**, 37–52.
8. Fisher, J. E., *Bot. Gaz.*, 1972, **133**, 78–85.
9. Rylski, I., in *CRC Handbook of Fruit Set and Development* (ed. Monselise, S. P.), CRC Press, Boca Raton, 1986, pp. 341–354.
10. Katoh, N. and Iwai, S., *Plant Cell Tissue Organ Cult.*, 1991, **26**, 7–10.

PIOUS THOMAS
J. B. MYTHILI

Tissue Culture Lab II,
Division of Biotechnology,
Indian Institute of Horticultural Research, Hessarghatta Lake,
Bangalore 560 089, India

Genotoxicity of 'gudakhu', a tobacco preparation used in Orissa

'Gudakhu' is a paste-like tobacco preparation the ingredients of which, besides tobacco (10%), are molasses (35%), lime (7%), red soil (a particular type of reddish soil available locally, 28%) and water (20%). During use it is rubbed over teeth with a finger tip. Its use is highly prevalent in Orissa and neighbouring states; in Orissa every third person is addicted to it. Persons are known using gudakhu for 40 years or more, and as many as 20 times a day.

Experimental and epidemiological studies have conclusively proven carcinogenic, mutagenic and clastogenic potential of tobacco and tobacco products[1–5]. Recently, a significant increase in the incidence of micronuclei has been recorded in exfoliated cells of buccal mucosa of habitual users of gudakhu[6]. Acetone extract of gudakhu (AEG) has also been proven to be genotoxic by us in mouse bone marrow system using chromosome aberration, micronucleus and sister chromatid exchange assays[7]. In view of the widespread and chronic use of gudakhu, it is worthwhile to study its genotoxic effect following repeated treatment. Bone marrow is a proliferative tissue and hence fails to provide correct information regarding genotoxic efficacy of an agent following chronic exposure. Chromosomal damages recorded in bone marrow cells at a particular sampling time following chronic exposure constitute the damages caused during the cell cycle prior to harvest, not the cumulative damages for the entire period of exposure. Sometimes an agent may go undetected in bone marrow cytogenetic assay following single treatment, particularly if it has low effect. Analysis of micronuclei (MN) in peripheral RBCs of mice, however, provides a very simple and important assay for evaluation of genotoxic efficiency of an agent following chronic exposure[8–10]. The present study deals with the evaluation of genotoxic effect of AEG in mice following repeated treatment for different periods using the micronucleus test (MNT) in peripheral erythrocytes.

Seven albino Swiss mice (4 males and 3 females) aged 10–12 weeks were employed. Gudakhu of 'Samaleswari' brand was purchased from the local market. Details of the extraction procedure have been presented elsewhere[7]. In brief, gudakhu of known weight was homogenized in a mortar and pestle with acetone, extract was filtered and the filtrate evaporated to dryness through vacuum desiccation. A freshly prepared dry pellet of known weight of gudakhu extract was suspended in an appropriate amount of distilled water and the suspension fed to mice by gavage. Animals were fed with a dose of 300 mg/kg (the highest dose tested in our earlier work[7] and having no visual abnormal effect on behaviour of the animals) of the dry extract once daily for 4 weeks. Blood was collected from the tail vein of each animal, without killing it, at the end of each week during the course of treatment and also on day 0 just before starting the treatment, and smear was drawn on grease-free slides. Blood collected on day 0 served as control sample. Thus, the same animals provided blood samples for control as well as for different test weeks of the treated series. Air-dried smear was fixed in absolute methanol for 15 min and stained next day in Giemsa diluted (1:10) with buffer (pH 6.8) for 10–15 min. Erythrocytes (both polychromatic and normochromatic) were examined for the presence of MN and the frequency of micronucleated erythrocytes (MNRBCs) was determined. For each sampling point 35,000 erythrocytes were scored from 7 animals (5000 from each).

Micronuclei were invariably round in shape and were of varied size groups (Figure 1). Never was an erythrocyte found to contain more than one MN. The frequency of MN increased significantly, compared to the control value, in all the test weeks. For weeks 3 and 4 the values were about 3 times the control value. Further, the increase showed a good positive correlation with the time course of chronic treatment (Figure 2).

The results clearly reveal positive genotoxic effect of AEG and support our earlier findings[6,7]. MN are induced by clastogenic effect or spindle dysfunction. The former generally results in induction of single micronucleus while the latter in multiple micronuclei. As in the present study, no case of multiple MN was encountered; the MN are supposed to have resulted from the clastogenic effect of AEG, and this finding is in accordance with our earlier one[7]. Elevated incidence of micronucleated polychromatic erythrocytes (MNPCEs) was also noted in peripheral blood of

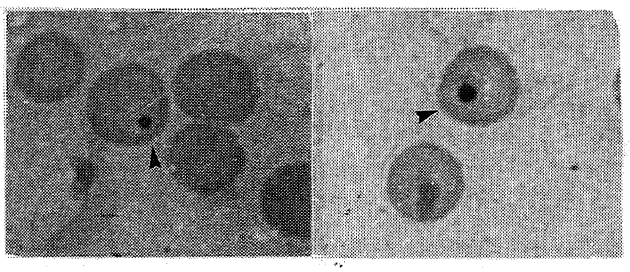


Figure 1. Micronucleated RBCs.

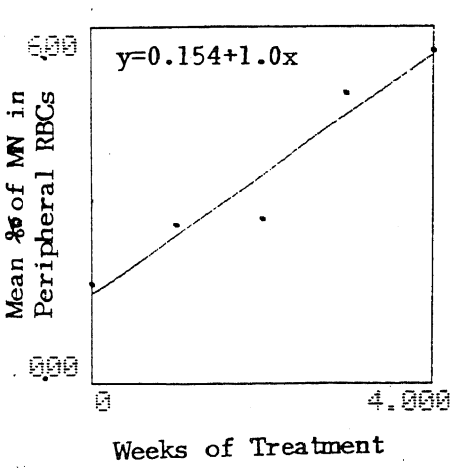


Figure 2. Linear regression curve showing the relation between the number of weeks of chronic treatment with gudakhu extract and the incidence of micronucleated peripheral erythrocytes ($r = +0.965$, $p < 0.001$, $a = 0.154$, $b = 1.00$).

new-born mice the mother of which had been exposed to tobacco smoke daily for 1 h during the last 6 days of pregnancy[11]. Recently, equal incidences of micronucleated peripheral reticulocytes (MNRETs) and bone marrow MNPCEs in mice treated with mitomycin C have been reported, though the induction of MNRETs was delayed by about 12 h compared to bone marrow MNPCEs[12].

In some strains of mice, unlike in most of the mammals, including humans, MNRBCs are not removed by the spleen; they remain in circulation[9,10,13,14]. As a result, MNRBCs get accumulated in their circulation. The RBCs of mice survive in the peripheral circulation for about 4 weeks[13]. Hence, repeated treatment is expected to exhibit gradual increase of the incidence of MNRBCs up to 4 weeks. Positive and significant correlation between the incidence of MNRBCs and the time course of chronic treatment obtained in the present study can be explained from such a behaviour of RBCs in this strain of mice.

Our results prove the uniqueness of this simple system[9,10], which is helpful in monitoring even weak clastogens following chronic exposure. Further, in this assay the same animals can provide blood samples, as in the present study, for control as well as treatment data for various periods. However, both polychromatic and normochromatic erythrocytes need to be considered.

Conclusion

Genotoxic effect of acetone extract of 'gudakhu', a paste-like tobacco preparation, has been evaluated using the micronucleus test in peripheral erythrocytes of mice following chronic exposure. A dose of 300 mg/kg of the dry extract in aqueous suspension was fed to mice once daily for 4 weeks. Micronucleated erythrocytes (MNRBCs) were scored from smears of tail vein blood collected just before beginning the treatment course and at the end of each week. The frequency of MNRBCs increased significantly in all the test weeks and showed a good positive correlation with the time course of treatment. The result has been discussed in the light of the fact that MNRBCs are not screened out by the spleen in mice, and this assay proved to be very important, particularly after chronic exposure.

1. Doll, R. and Peto, R., *J. Natl. Cancer Inst.*, 1981, **66**, 1191–1308.
2. *Tobacco Habits Other Than Smoking*, Monograph, no. 37, IARC (International Agency for Research on Cancer), Lyon, 1985, pp. 113–255.
3. *Tobacco Smoking*, Monograph no. 38, IARC (International Agency for Research on Cancer), Lyon, 1986, pp. 60–298.
4. Kukarni, J. R., Sarkar, S. and Bhide, S. V., *Mutagenesis*, 1987, **2**, 263–266.
5. Stich, H. F. and Anders, F., *Mutat. Res.*, 1989, **214**, 47–61.
6. Das, R. K. and Dash, B. C., *Fd. Chem. Toxicol.*, 1992, **30**, 1045–1049.
7. Dash, B. C. and Das, R. K., *Mutat. Res.*, 1992, **280**, 45–53.
8. MacGregor, J. T., Wehr, C. M. and Gould, D. H., *Environ. Mutagen.*, 1980, **2**, 509–514.
9. Schlegel, R. and MacGregor, J. T., *Mutat. Res.*, 1982, **104**, 367–369.
10. Schlegel, R. and MacGregor, J. T., *Mutat. Res.*, 1983, **113**, 481–487.
11. Balansky, R. M. and Blagoeva, P. M., *Mutat. Res.*, 1989, **223**, 1–6.
12. Hayashi, M., Morita, T., Kodama, Y., Sofuni, T. and Ishidate, M., Jr., *Mutat. Res.*, 1990, **245**, 245–249.
13. Mavournin, K. H., Blakey, D. H., Cimino, M. C., Salamone, M. F. and Heddle, J. A., *Mutat. Res.*, 1990, **239**, 29–80.
14. Heddle, J. A., Cimino, M. C., Hayashi, M., Romagna, F., Shelby, M. D., Tucker, J. D., Vanparys, P. and MacGregor, J. T., *Environ. Mol. Mutagen.*, 1991, **18**, 277–291.

ACKNOWLEDGEMENT. This work was supported by the UGC Grant no. F3-137/87 SR–II sanctioned to RKD.

R. K. DAS
B. C. DASH

School of Life Sciences,
Sambalpur University,
Jyotivihar 768 019, India

Azolla sp. from the Early Cretaceous of Cauvery basin, South India

A new record of *Azolla* sp. based on megaspores and massulae from the Late Albian (Early Creataceous) of the Pondicherry area in South India is presented in this study. The massulae are with characteristically circinate and septate glochidia. This is the oldest record of *Azolla* known to date.

During a palynological study of the Early Cretaceous carbonaceous shales from a depth of 32.4 m in bore hole BH/23, drilled by a private party for ground water, 1.3 km west of Kattarambakkam village in Pondicherry area in South India, the authors have come across a number of well-preserved massulae of *Azolla*, a water fern, along with varied spore and pollen taxa. One of the palynoslides also shows fortuitously a fairly large megaspore with a thick mat of hairy (tomentose) perine in which are seen entangled five massulae (Figure 1 *a*). All these massulae are characterized by circinate glochidia. Similar massulae have not been recorded to date from any horizon in India.

The following are the details of this new find of *Azolla* sp.: (i) megaspores about 500 μm in diameter, extremely thick-walled, wall up to 30 μm thick; stratification not clear as the specimen is dark, dense and more or less opaque; (ii) outer layer of megaspore wall representing perine well developed and made up of thick felt of delicate hairs (tomentose) (Figure 1 *a–c*); (iii) number and nature of floats could not be deciphered; (iv) massulae attached to perisporal hairs of megaspore, oval to elliptical or irregular in shape, 85–214 μm in longest dimension, 54–125 μm in shortest dimension; (v)

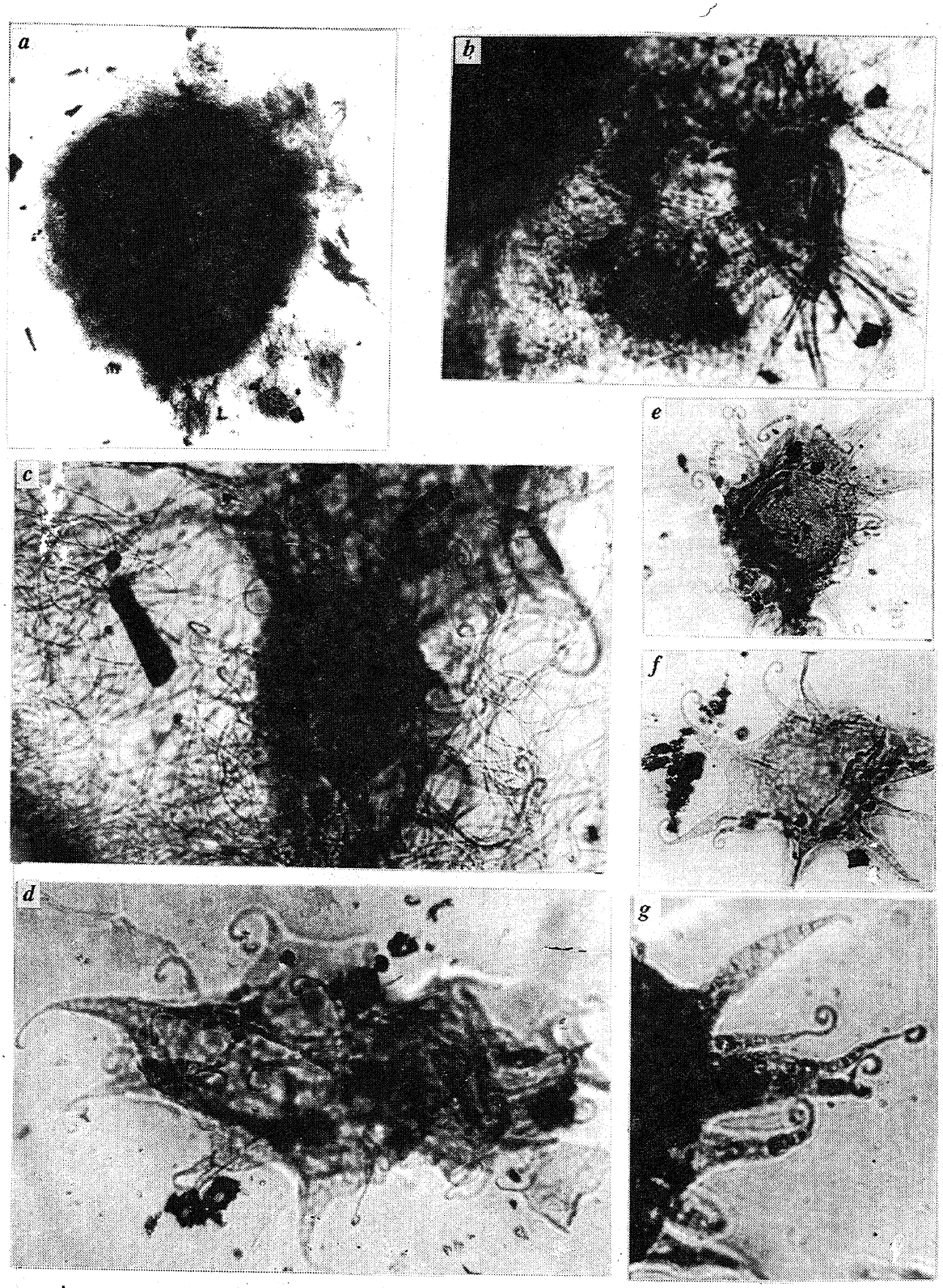


Figure 1. *a*, *Azolla* sp. Megaspore surrounded by massulae (×100). *b*, *c*, Part of a megaspore (enlarged) showing perisporal hairs with massulae entangled (×300). *d*, Massula with circinate, septate glochidia and a microspore (×400). *e*, *f*, Smaller massulae (×300). *g*, Glochidia enlarged (×500).

surface alveolar and foamy, (vi) glochidia many, 7–10 μm broad at base, 42–60 μm long, filamentous, mostly circinate at apex, usually with one gyre, occasionally with only a partial gyre, distinctly septate (Figure 1 *d–g*); (vii) basal part of glochidion often robust; (viii) microspores 1–3 per massula; 32–50 μm in diameter (few up to 60 μm), thick-walled, psilate to locally granular.

The spore and pollen complex recovered from the samples which contained *Azolla* remains is represented by the following significant taxa: *Lycopodiumsporites austroclavatidites*, *Ceratopteris equalis*, *Dictyosporites speciosa*, *Crybelosporites stylosus*, *Cicatricosisporites australiensis*, *Plicifera* sp., *Coronatispora perforata*, *Contignisporites glebulentus* sp., *Appendicisporites tricornitatus*, *A. distocarinatus*, *Schizosporis reticulatus*, *Triporoletes reticulatus*, *T. radiatus* (common), *Tricolpites* (few species), etc.

This assemblage is strikingly similar to the palynoflora of the Late Albian deposits of the Cauvery basin[1]. The age of the *Azolla*-bearing deposits under study is, therefore, assigned to Late Albian (Early Cretaceous). Hitherto the oldest records of *Azolla* (*A. circinata*, *A. montana* and *A. geneseana*) were from the Maestrichtian (Late Cretaceous) of Montana[2], USA, and Alberta[3], Canada. Though resembling the present specimens of *Azolla* in the possession of circinate glochidia, the American *A. Circinata* is distinguishable in its non-septate and much shorter glochidia. The present record of *Azolla* thus constitutes the oldest and the massulae-bearing circinate and septate glochidia probably represent the primeval condition. A detailed study involving the nomenclature and evolutionary aspects of the new record of *Azolla* sp. is in progress.

1. Venkatachala, B. S., Rawat, M. S., Sharma, K. D. and Jain, A. K., IV Int. Palynol. Conf. Lucknow (1976–77), 1980, **2**, 396–402.
2. Hall, John W., *Am. Fern J.*, 1968, **58**, 77–88.
3. Hall, John W. and Swanson, N., *Am. J. Bot.*, 1968, **55**, 1055–1061.

ACKNOWLEDGEMENTS. We are thankful to Mr D. C. Das Sarma, Director, Palaeontology Lab., GSI, Southern Region, Hyderabad, for constant encouragement while carrying out the studies. We thank P. Kumaraguru and Mr Kanishkhan from GSI, Madras, for the borewell samples.

R. S. PATIL*
M. RAMA RAO*
C. G. K. RAMANUJAM**

**Palaeontology Division, Geological Survey of India, Southern Region, Hyderabad 500 068, India*
***Department of Botany, P. G. College of Science, Saifabad, Hyderabad 500 004, India*

Angioarchitecture of the glomerulus of a swamp catfish, *Chaca chaca* (Ham.) (Chacidae, Siluriformes)

Glomerular vascular architecture of a freshwater swamp catfish, *Chaca chaca* (Ham.) was investigated using corrosion vascular resin replica, a novel method to study angioarchitecture of an organ at capillary level. Each afferent glomerular arteriole (AGA) ramifies to form tortuous capillaries of almost round glomerulus. The tortuous glomerular capillaries converge to form efferent glomerular arteriole (EGA) of almost similar dimension.

The past light and ultrastructure studies have revealed fenestrated glomerular endothelium[1], the nature of podocytes[2] and the variation in the thickness of glomerular capillary in freshwater and marine teleosts[3]. However, little is known on the microvasculature of the glomerulus of the teleostean species[4]. Paucity of information on glomerular vasculature forms the basis of the present communication on the angioarchitecture of the glomerulus of a swamp catfish, *Chaca chaca*.

Chaca chaca is a curious-looking catfish inhabiting the swamps of northern Bihar (India) infested with macrovegetation and decaying organic matters. This fish is well adapted to such hypoxic swamps. Live specimens of *C. chaca* were collected with the help of local fishermen from the swamps near Purnia (Northern Bihar, India) and transported to and maintained in the glass aquaria (40 l) of the department.

The fish (65 g) was anaesthetized in MS 222 (ethyl-m-aminobenzoate). The fish was carefully operated and its ventral aorta was cannulated via the exposed bulbus aorta. The entire vasculature was flushed with heparinized phosphate-buffered Ringers solution. Methyl methacrylate (Mercox) was infused at 25–30 mmHg pressure until the resin polymerized into plastic[5]. The mercox-perfused fish was immersed in warm (50°C) water for about 1 h to ensure complete polymerization. The fish was treated with 20% NaOH, distilled water and HNO_3 to obtain tissue-free vascular replica. A part of renal vascular replica was processed and examined under P SEM-500 at RSIC, Bose Institute, Calcutta.

From the corrosion vascular replica it is evident that the renal arteries of the two sides of kidney originate alternately from the dorsal aorta at an interval of about 0.93 mm. The renal artery runs away from the dorsal aorta transversely for about 233 μm and then turns at right angle, runs for about 700 μm to give rise to the glomerular artery (Figure 1). The glomerular artery connects the tortuous glomerular capillaries through AGA. The tortuous capillaries are perhaps associated with lowering the velocity of the blood flowing through the glomerulus for its effective renal functioning. In narrow capillaries (diameter about 7 μm) red cells flow in their centre in a single file. Such a flow pattern is designated as axial train type. In wider capillaries (diameter about

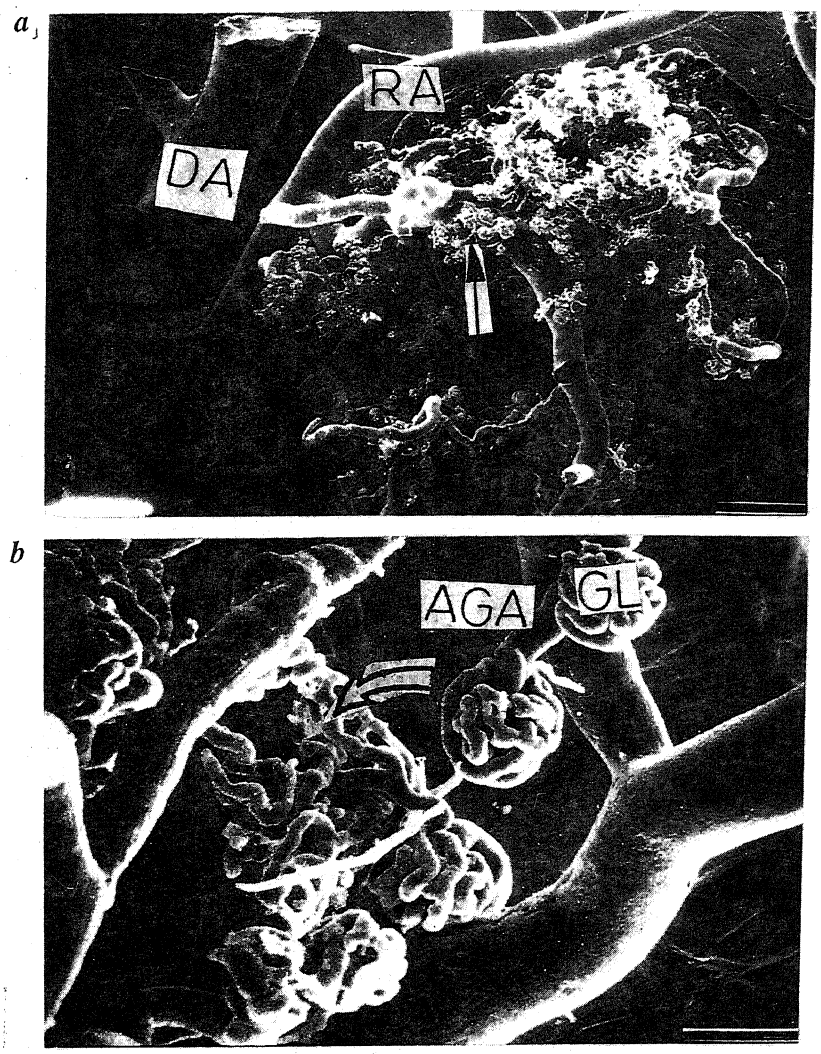


Figure 1. *a*, Corrosion vascular replica of a part of renal tissue showing dorsal aorta (DA), glomeruli (arrow), glomerular artery (arrowhead), and renal artery (RA), bar = 700 µm. *b*, Magnified view of above showing afferent and efferent glomerular arterioles (AGA, EGA), arteriovenous anastomoses (bent arrow), glomeruli (GL) and vein (V), bar = 50 µm.

11 µm) red cells flow in a multiple file. This type is referred to as shearing core type. When the capillary diameter varies between 8 and 10 µm, both types of flow patterns can occur[6]. Narrow capillaries (diameter 7.08 ± 1.2 µm) and the dimension of the short axis of red cells (5.6 µm) suggest[7] single axial file of erythrocytes flowing with peripheral plasma layer in the glomerular capillaries of *Chaca*. Like carp (*Cyprinus carpio*) and unlike marine fish[8], the AGA and EGA are of similar dimensions to ensure comparatively higher filtration rates at lower pressures in the glomerulus. To the contrary, wider AGA and narrow EGA of marine fish are associated with low filtration rates at higher pressures[8]. The EGA ramifies into arteriovenous anastomoses before its blood is drained into the renal vein.

1. De Ruiter, A. J. H., *Cell Tissue Res.*, 1980, **206**, 1–20.
2. Brown, J. A., Taylor, S. M. and Gray, C. J., *Cell Tissue Res.*, 1983, **230**, 205–218.
3. Hickman, C. P. and Trump, B. F., in *Fish Physiology* (I) (eds Hoar, W. S. and Randall, D. J.), Academic Press, New York, 1969, pp. 91–239.
4. Endo, M., *Jpn. J. Ichthyol.*, 1989, **36**, 346–349.
5. Olson, K. R., Munshi, J. S. D., Ghosh, T. K. and Ojha, J., *J. Morphol.*, 1990, **203**, 165–179.
6. Whitmore, R. L., *J. Appl. Physiol.*, 1967, **22**, 767–771.
7. Mishra, N., Towheed, M. A., Das, A. K., Singh, B. N. and Ojha, J., *Proc. Indian Natl. Sci. Acad.*, 1992, **B58**, 341–346.
8. Guyton, A. C., in *Text Book of Medical Physiology*, W. B. Saunders, Philadelphia, 1986, 7th edn, pp. 358–473.

ACKNOWLEDGEMENT. Financial assistance from UGC, Department of Environment, New Delhi, and SEM facilities from RSIC, Bose Institute, Calcutta, are gratefully acknowledged.

J. OJHA

Postgraduate Department of Zoology,
Bhagalpur University,
Bhagalpur 812 007, India

Location and magnitude of the 1833 Nepal earthquake and its relation to the rupture zones of contiguous great Himalayan earthquakes

Roger Bilham

Accounts of an earthquake on 26 August 1833 which was felt over a large part of northern India have been interpreted by some authors to represent a great Himalayan thrust event beneath Western Nepal. However, details of the event in the Indian press of 1833 and scientific journals of that time suggest that the epicentre of the earthquake was near Kathmandu, within or close to the inferred rupture zone of the 1934 Bihar earthquake. Estimates of the moment magnitude based on reported intensities indicate that the earthquake was of $7.5 < M < 7.9$, and as such may have done little to release the elastic strain accumulating in the region of the Central Himalayan Seismic Gap, contrary to the expectation of some authors. The location of the epicentre was probably N or NE of Kathmandu, adjoining or overlapping the rupture area of the great 1934 Bihar/Nepal earthquake. The moment magnitudes of great Himalayan earthquakes in 1897, 1905, 1934 and 1950, and of smaller recent events, are compared using recently published empirical relations between isoseismal areas and moment magnitude. When due allowance is made for deficiencies in field data, reasonable fits are obtained for all events except for the 1905 Kangra earthquake. The intensity VIII area for this event is anomalously small for an $M \geq 8$ earthquake associated with several meters of slip. It is proposed that the Kangra earthquake may have been a slow earthquake.

APPROXIMATELY half of the Himalayan arc has ruptured in four great earthquakes in the past 100 years (Figure 1). The largest region between the rupture zones of these recent events is a 500–800 km segment of the Himalaya between the 1905 Kangra and the 1934 Bihar earthquakes, approximately between the longitudes of Kathmandu and Delhi. Of importance in estimating the present slip potential of this segment, termed the *Central Gap* by Khattri and Tyagi[1], is the existence and severity of great historic earthquakes that may have ruptured all or part of the gap. A severe earthquake occurred in Nepal in 1255 when 'innumerable towns were utterly destroyed and thousands of their inhabitants killed'[2] but the regional extent of this event is unknown. Other large pre-XX century earthquakes in Nepal (1408, 1681, 1810, 1833 and 1866) are mentioned by Chitrakar and Pandey[3] but none appear to have been as damaging as the 13th century event, causing concern that considerable elastic strain may be available presently to drive one or several $M > 8$ earthquakes in the Central Gap.

An alternative mechanism to absorb slip between Tibet and India is to invoke the possibility of aseismic slip (slow earthquakes or creep) over at least a part of the region. Levelling data and recent GPS measurements between India and central Nepal[4–6] suggest that creep processes that might otherwise release Indo-Asian convergence aseismically have been insignificant in the past few years. If similar creep rates (2.5 ± 2.5 mm/yr) exist elsewhere along the arc throughout the seismic cycle, they are evidently inadequate to accommodate completely the slip budget between India and southern Tibet, although they may delay rupture[7]. The possibility that some Himalayan earthquakes may be slow events, with large slip but little radiated high-frequency seismic energy, cannot be excluded[8,9]. Such events would not appear in the historical record as great earthquakes although they could, in principle, release the elastic strain associated with one.

Earthquakes in 1803, 1833 and 1866 appear to have occurred at least partly within the Central Gap[10] and the largest of these in terms of felt area is believed to be the 1833 event. Reports of the 1833 earthquake are found in newspapers starting the day after the earthquake, and these and other data are collated in three issues of the *Journal of the Asiatic Society of Bengal* in the months following the earthquake by Prinsep[11] and Campbell[2,12], and by Baird-Smith[13,14] in two articles a decade later. Summarises of these summaries are found subsequently in various catalogues and comparative studies[15–21]. The compilation by Dunbar *et al.*[21] lists the event as severe

Roger Bilham is at CIRES and the Dept. of Geol. Sci., University of Colorado, Boulder CO, 80309-0216, USA.

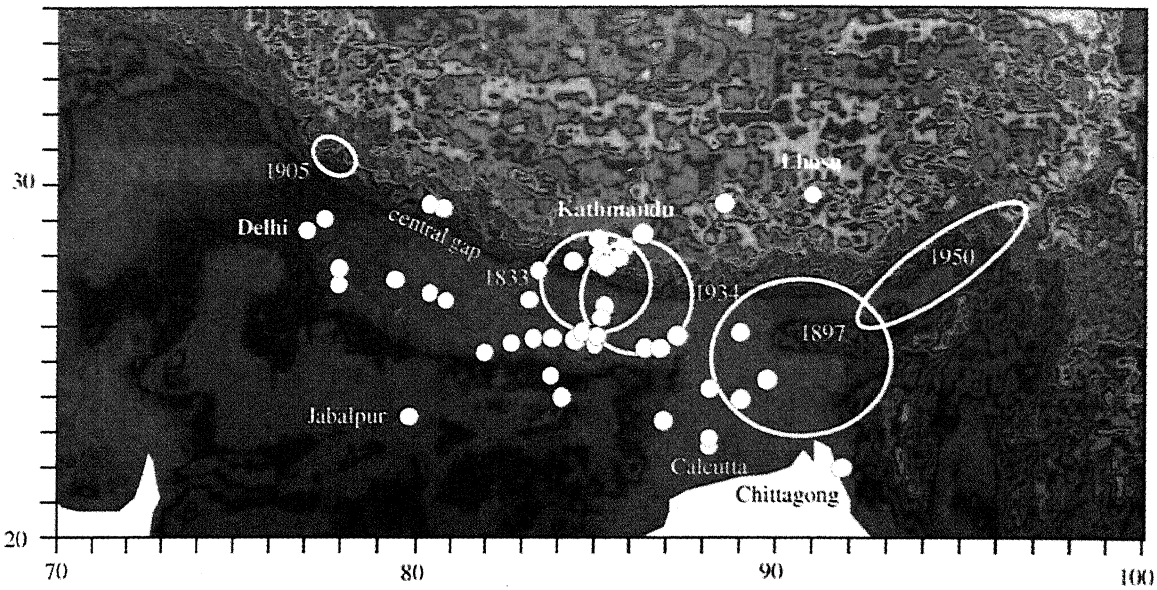


Figure 1. Approximate Modified Mercalli Intensity VIII zones for 4 great earthquakes in the Himalaya since 1833 and settlements reporting information concerning the 1833 shock (white dots). A great earthquake may not have occurred in the 800 km segment of the Himalaya between the 1905 Kangra and 1934 Bihar earthquakes (Central Gap) for 650 years.

and records its location as 25.1°N and 85.3°E, near Patna, south of the River Ganges, at the southern limit of intense shaking described in 1833 reports. A location west of Kathmandu is favoured by some authors[22], who suggest tentatively that it may have occurred in the Central Himalayan Gap. Khattri and Tyagi[1] place the earthquake approximately 130 km west of the Bihar 1934 epicentre on the edges of the Central Gap and assign the event $M = 7.6$, a location and magnitude consistent with the findings of the present study. One purpose of this article is to estimate more precisely the location and magnitude of the 1833 Nepal earthquake using authentic accounts found in newspapers and scientific articles published soon after its occurrence. A second purpose of the article is to reconcile the intensity data from the 1833 earthquake with the felt reports of four great Himalayan earthquakes and other earthquakes that have occurred since then.

The 1833 earthquake

Just before midnight on 26 August 1833 (23:35 Calcutta time), a 1 million km^2 region of northern India, Nepal and Tibet was shaken by a strong earthquake which triggered landslides and rockfalls, destroyed more than 4600 dwellings and many temples, but apparently resulted in fewer than 500 fatalities. It is certain that the loss of life would have been far more severe had not the main shock been preceded by two large foreshocks 5 h and 15 min before the main shock that drove people outdoors in alarm. In some villages in northern Nepal and southern Tibet, 30% or more houses were destroyed. Damage in India was less severe with fewer than 10 reported fatalities and few buildings totally destroyed.

The main shock was felt from Delhi to Chittagong (Figure 1). It was not felt in Lhasa, nor are there reports of the shock further SSW than Jabalpur. Intensities were high in the mountainous region near and north of Kathmandu, and reduced rapidly to the south and more slowly to the east and west. Intensities remain high at the northern limit of high-intensity observations in southern Tibet. Accounts of damage where shaking was most intense suggest a similar intensity distribution to that observed during the Bihar 1934 earthquake with the principal exception that accounts of liquefaction features in 1833 are rare. Passes to Tibet were blocked by landslides, and the Kamla River was dammed by a landslide that burst 4 days after the event flooding the village of Baldeah near the India/Nepal border (Bengal Hurkaru, 16 September 1833).

Isoseismal coverage and accuracy

The fidelity with which the Indian newspapers reported the effects of the 1833 earthquake can be gauged by the faithful reprinting of articles printed in one by another. More than a dozen newspapers and weekly and monthly journals were published in Calcutta (The *Reformer*, Calcutta, **3**, 141, Oct. 27, 1833) although not all of them reported the news of the event, and a few are no longer available. Newspapers in Delhi, Bombay, Meerut and Madras also contain accounts of the earthquake. In the 52 accounts in the Appendix, 97 towns and villages are mentioned. Most of these reports can be assigned Modified Mercalli Intensities V–IX with reasonable certainty. Few 'barely felt' reports are available and fewer 'not felt' reports (e.g. Nazirabad, Xigatse and Lhasa), and the general sparcity of intensity IV or smaller observations prevents reliable estimates of those areas that ex-

perienced the earthquake weakly. The absence of details of construction methods of the houses destroyed, or of those surviving, means that intensity estimates are not as reliable as those for the 1934 earthquake[19], and although it is likely that construction practices changed little between the two earthquakes, it is certain that urban dwelling units and civic buildings were less numerous in 1833. This may influence comparisons between the 1833 and 1934 earthquakes because for both the earthquakes several damage reports are from urban regions near factories or administrative centres. The Gangetic Plain except for the Kingdom of Oudh, south of western Nepal, was under British colonial rule in 1833 (ref. 23). However, reports from Oudh are available despite a significant cholera epidemic in Lucknow and Allahabad. With perhaps the single exception of data from Monghyr (Monger) most of the reports contributed to the Indian newspapers are in the form of personal accounts of the earthquake with no statistical descriptions of urban damage. Few fatalities are reported: three during the main shock at Chapra and one during an aftershock in Bhagalpur. In some accounts additional fatalities are implied although it is probable that the low loss of life was real, and is attributable to the second foreshock that evidently drew people into the safety of the open air fewer than 15 min before the mainshock.

In contrast to the anonymous accounts in the Indian newspapers, Archibald Campbell, an assistant surgeon to the British Residency in Kathmandu, in three contributions to the Asiatic Society of Bengal[2,11,12] compiled unofficial numerical damage and casualty data from the Kathmandu Valley, from travellers returning along the Kodari route from southern Tibet (passing through Dulka, the northern Nepalese border district of Dolakha), and from the official delegation returning from Lhasa through southern Tibet. He also mentions damage to forts along the southern trading route from India (Hetauda, Chisapani). Although he also obtained information from western and east-central Nepal, his sources of information from these regions are not cited and data from these regions may be less reliable. Thus, unlike data from N and NE Nepal, which mention specific towns and villages, his information from western Nepal refers to regional administrative areas: 'At Gorkha (75 km E of Kathmandu), only two houses were destroyed; at Palpa (180 km E of Kathmandu) none; and at Doti (540 km E of Kathmandu in easternmost Nepal), on the borders of Kumaon, the shock was felt, but by no means severely.' Although this westward decay in intensity over several hundred km is unsupported by details, it is consistent with the decay in intensity documented independently in northern India. For example, Baird-Smith[13] reports that 'it was scarcely felt at all' at Lohughat (=Lohu Ghar, 29°23′N, 80°05′E, 18 km from the E. Nepal border). Campbell[2,12] also emphasizes that the shock was not felt in Lhasa nor Xigatse, contrary to Mallet[16] and Dunn *et al.*[19], who repeat early speculation by Prinsep[11]. In a following issue, Prinsep, as editor of the Asiatic Society journal, corrected his erroneous guess that Lhasa was close to the epicentre[2,12] but not until it had been repeated in Calcutta newspapers[24].

In general terms, information concerning the 1833 earthquake in northern India is restricted to British trading posts, and information from Nepal and Tibet appears to be restricted to the Kathmandu Valley and to villages on merchant routes from Kathmandu to Tibet and to India (Figure 3 and Tables 1 and 2).

Location names

Place names in India are readily matched with current locations for all but three 1833 listed localities: Tirhoot, Mullye (Mallai of Prinsep[11]) and Baldeah. The first of these corresponds to the district of Tirhoot south of Nepal named after the town of that name, now Muzaffarpur (District Map of India, 1840, 64 mi = 1"); however, it is not clear that the three damage reports from Tirhoot came from Muzaffarpur or from villages between Muzaffarpur and the Nepal border (Motihari, Sitamarhi, Darbhanga, etc.). Despite the absence of specific locations, the high-intensity shaking reported from Tirhoot unquestionably identifies a region in the southern Terai of Nepal or northern plains of India.

Mullye has eluded identification with either a district or town, and in this case the absence of a precise match is regrettable because from this region comes one of the two comparisons of the 1833 event with the 1803 Kumaun earthquake (the other account from a resident in Calcutta recalling the occurrence of a severe earthquake during the siege of Aligarh (27°55′N, 78°10′E) soon after the start of the second Anglo/Maratta war). The approximate location of Mullye is revealed by the fact that Calcutta news reports from Mullye took 9–11 days between mailing and publication. An additional clue is that a summary of the effects of the event south of 26°N from a correspondent in Agra omits mention of Mullye, suggesting that it lies to the north of 26°N. Mallet[15], who in 1851 discusses the 1833 earthquake as the 1834 Nepal earthquake places 'Mallai?' on a travel time map of India (*ibid.*, Plate XVII) at approximately 22.9°N, 91.2°E (Maijdi); yet the inferred intensities are too high, and the reported sun-clock times of the main shock too early, for this location. He corrects the date to 1833 in his report[16] of 1854. A graph of news travel times to Calcutta (Figure 2) shows that an eastward location for Mullye could correspond to the Vedic district of Malla, approximately between Gorackpur and Chapra, but this term is unlikely to have been in common use in the 19th century. Of towns and trading centres in the region, Mullye is considered a possible abbreviation for Malihabad or Malloon[23] near Lucknow,

Manjhi (25°40′N, 84°30′E) near Chapra, or Murliganj (25°42′N, 86°59′E) near Purnia. Malihabad is evidently too far from Calcutta for normal 9 day news delivery, requiring travel times exceeding 120 km/day. Although the remaining suggested locations are at the correct range for news delays of 9–11 days, the phonetic resemblances are poor. In the absence of certain identification the three intensity V reports from Mullye are omitted from the isoseismal map.

It has not been possible to identify the precise location of Baldeah on the Kamla River, where a flash flood followed the breaching of a temporary lake dammed by an earthquake-triggered landslide. From the description (Appendix) it is likely that this was near the Nepal border, perhaps at Batia 27°15'N, 86°08'E.

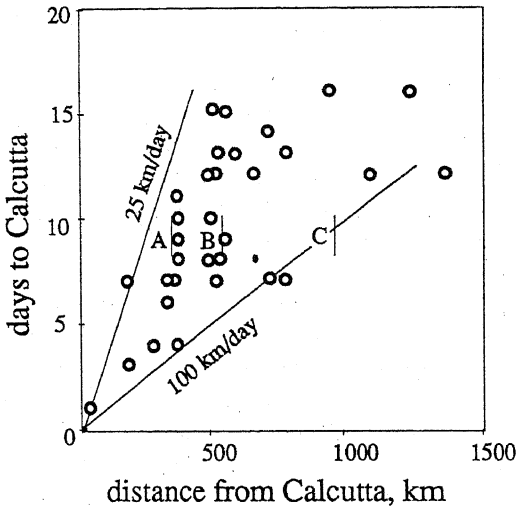


Figure 2. Delay between mailing a letter and its arrival in Calcutta assuming that publication occurs the day following its arrival. The three vertical lines indicate the possible locations for Mullye: A = Murliganj; B = Manjhi; C = Malihabad or Malloon. The Vedic district of Malla lies between B and C. Distances are geodesic. The most remote point on the graph from Calcutta is Delhi.

Place names with phonetic similarity and geographic proximity to the indicated regions have been identified for most of the villages in the Kathmandu Valley, and where these are questionable they have been identified in Table 1. Although relatively few locations outside the Kathmandu Valley are mentioned by Campbell, they are important in identifying shaking intensity to the north and east.

Campbell's latitude for Tingri (28°N) is approximately ≈ 60 km too far south, suggesting that he did not have good maps of southern Tibet. Rennell's 1782 Map of India (Figure 3), showing Tingri at 29.3°N, was known to be in error to the British Residency in Kathmandu in 1801 but new maps prepared by Crawford in 1803 of the route to Dugurcheh (Xigatze) were very incomplete[25]. Rennell's map is of interest because it shows Nesty, Dunna, Kansa, Chuska, Kut or Kutí, and Mescingzhung in sequence northward toward Tingri, over the mountain pass between Nepal and Tibet near the present border village of Kodari. The spacing of villages suggests that each interval represents a day's march. Nesty appears to correspond to Nisti, and the old fort at Dugunna Garhi may correspond to Dunna. Kansa is the village of Dram (Khasi in Tibet), 5 km NE and 400 m above Kodari, corresponding to Kassa mentioned by Campbell[12] without explicit damage reports. Chuska of Rennell's map may correspond to Choksum (28.07°E, 86.00°N) or Kan Sing Chok[2,12]. The location of Kuti mentioned several times by Campbell is the Tibetan name for Nyalam (D. Bresheares, personal communication 1994). At least two of the five villages mentioned as on the 'Bhote or eastern pass to Tibet' (Nan Sing Chock, Kuti, Kassa, Mundun Pahar and Listigaon) appear to correspond to villages in a 20 km NS region of intense shaking that includes the present border village of Kodari, and it appears possible that information from a region within a quarter of a degree of 27°N, 28°E was conveyed by travellers returning from Tingri to Kathmandu along this trading route.

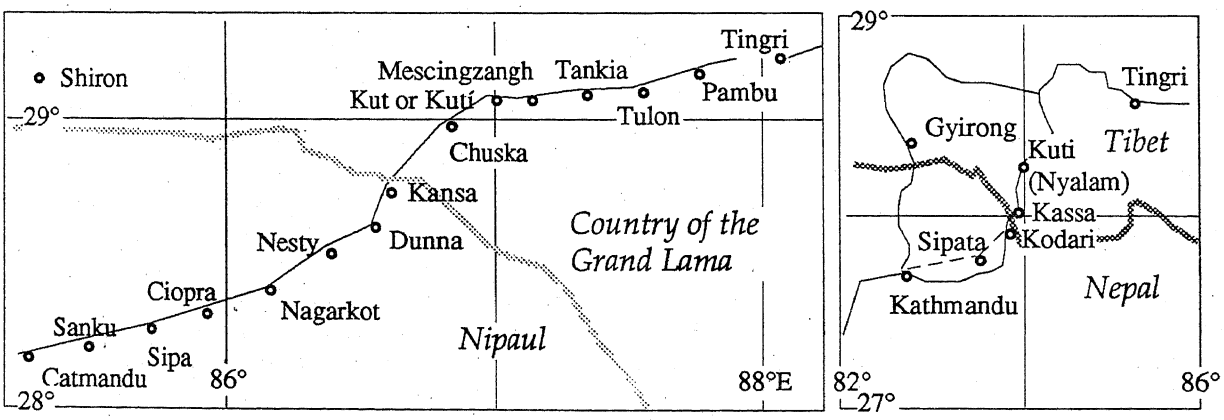


Figure 3. Villages between Kathmandu and Tingri through the Kuti Pass (Rennell's 1782 Map of India). Figure on the right shows the true coordinate locations for Kathmandu, Tingri, Gyirong, inferred locations for Kassa, Sipa, Kuti, and merchant routes to India and Tibet.

Rennell's map also shows the village of Sipa (Campbell's village of Shipa, 32 km E of Kathmandu) on the Kofs (Kosi) River between Sanku (Sankhu, 27°45′, 85°28′) and Ciopra (?Chautara). Sipa probably corresponds to Sipata or Sipatinghare (27°45′, 85°39′) on the Indrawatty tributary of the Sun Kosi, which lies on the trekking trail to Kodari north of the current Kathmandu–Kodari road. The village of Nagarkot (meaning Snake Hill) on Rennell's map between Ciopra and Nesty does not correspond to the geodetic observatory of that name on the eastern hills overlooking the Kathmandu valley.

The location of places in Campbell's account[2] are of particular importance in tracing the felt area to the north and east of Kathmandu. Digarchi is the present Shigatse or Xigatze (29°17′, 88°54′) and Tingri (28°30′, 86°30′) retains its old name. However, Campbell indicates that from Tingri a 8–10 day march due west brought the Nepal/Beijing returning delegation to the village of Kirung. Kyirong or Gyirong (28°33′, 85°16′, Shiron on Rennell's map) is 180 km W along the road from Tingri, and although the pass south of Kyirong enters Nepal by the Bhote Kosi, which means 'Tibet sacred river', it is only one of several rivers of that name crossing the border, including one through the Kuti pass into Kodari. It is for this reason that Campbell may have been misled into entering in parenthesis that the delegation came through 'the eastern pass of the valley into Bhote'. The Kyirong route they followed would more appropriately be called the northern pass to Tibet.

Isoseismal map for 1833

Descriptions from the original reports were assigned values on the Modified Mercalli Intensity scale (Tables 1 and 2 and Appendix). Where several reports from the same village exist, the intensity data in most cases agree, adding confidence to the observations. Moreover, although the intensity data for the 1833 event are somewhat uneven in coverage, they yield relatively smooth spatial variations in intensity (Figure 4). The intensity data may be interpreted as smooth curves that approximately encircle the appropriate intensity data, or more complex figures that adhere rigorously to mapped intensities. Figure 4 shows a series of smooth ellipses that approximate the observations. The resulting isoseismals are prejudiced by a preference for simplistic geometry; yet, though a number of alternative curves could be constructed, they would have the unavoidable characteristic of indicating maximum intensities in east-central Nepal, close to the 1934 epicentral region. A more realistic set of isoseismal curves is considered unwarranted given the distribution of intensity data. As an example of the complexity denied by adopting elliptic isoseismals, the intensity observations from the 1833 earthquake are superimposed on a map of the observed and inferred isoseismals for the 1934 Bihar earthquake. Few of the 1833 intensity data are found outside the corresponding 1934 isoseismals, suggesting that similar complexity in local isoseismal geometry may have prevailed in the two events, enforced by basin response characteristics.

An approximate estimate for the felt area of the 1833 earthquake (dashed on Figure 4) is found in Oldham's memoir on the 1897 earthquake[18]. Oldham does not indicate his sources although it is likely that the data came from his father's chronology of Indian earthquakes[12] and that the contour shown is meant to represent the area within which the earthquake was perceptible (i.e. Mercalli intensity II). This 0.21×10^6 km^2 felt area includes Chittagong to the southeast (editorial footnote by Prinsep[11] 'Mr. Walters informs me that it was also felt in Chittagong'), but omits information from Delhi and other western points, and from Jabalpur to the SW. The use of a smooth ellipse to embrace these western points increases the felt area to approximately 1.2×10^6 km^2. Uncertainty in this estimate is caused largely by the sparcity of data to the SW, E and N. Although the elliptical aspect ratio of the intensity $\geq$ VI isoseismals is favoured by the specific absence of felt shaking in Lhasa and Xigatse and by the shape of intensity VII isoseismals, the absence of reports from central India in theory permits the felt radius to increase substantially to the SW. Thus, the estimated areas for intensity $<$ V are assumed to err on the low side and estimates for earthquake magnitude presented below are given both for elliptical and circular isoseismals. Intensity I–III isoseismals are omitted from Figure 4 as no data are available. The largest ellipse drawn for the 1833 earthquake in Figure 4 corresponds to Mercalli Intensity IV, which is presumed to be significantly smaller than the area of felt perceptibility. Reports from locations specifying that the earthquake was not felt are absent except for the NE, and from remote locations such as Bombay and Madras.

Uncertainties in isoseismal area in Table 3 correspond to maximum and minimum estimates for areas contained within the smooth symmetrical contours selected to fit the observations in Figure 4. Although these are apparently uncertain to ±1% based on the mapped isoseismal reports, this figure does not account for errors in estimating the Mercalli Intensity from the eyewitness accounts reported in the Appendix. Variations in building toughness and subjective interpretation of damage may result in errors perhaps an order of magnitude larger. Moreover, from the 1934 (ref. 19) and 1988 (ref. 26) Bihar/Nepal events it is clear that isoseismal contours in the region are typically more complex.

The congruence between the 1833 and the 1934 intensity data is interesting in that in the few locations where isoseismic intensities disagree they differ by not

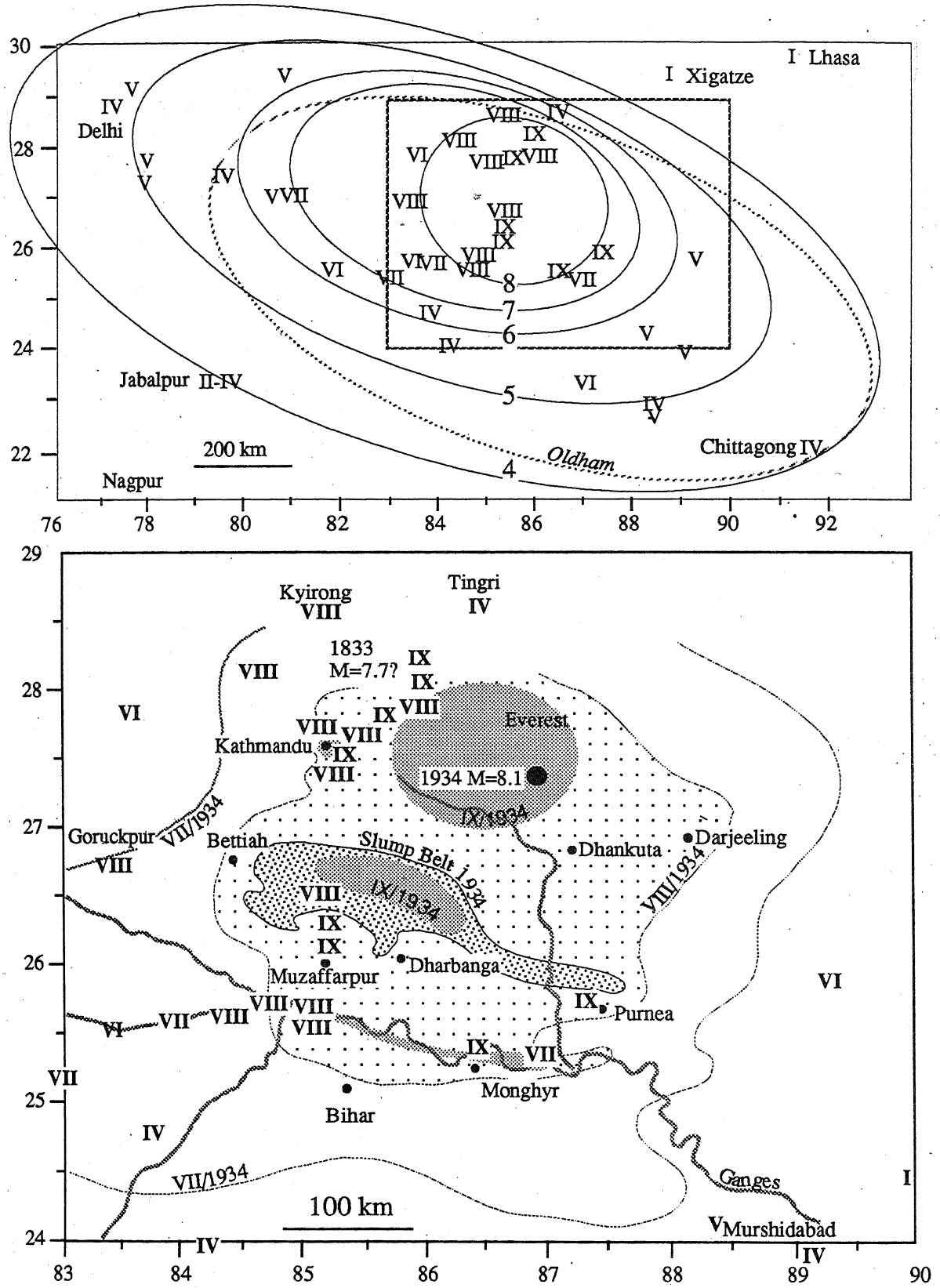


Figure 4. Intensity reports for the 1833 Nepal earthquake with smooth elliptical isoseismals. In the enlarged view of the epicentral region (inset above) the 1833 intensities are superimposed on isoseismals VII–IX for the Bihar 1934 earthquake (modified from Dunn *et al.*[19] by incorporating the data discussed by Pandey and Molnar[27]. On both occasions high-intensity shaking occurred in the same geographic settings although the 1934 isoseismals appear displaced approximately 100 km to the east. A possible 1833 epicentral location and moment magnitude is indicated.

Table 1. Summary of damage in Nepal and Tibet[2,11,12]. 414 fatalities, 172 injuries and 4040 destroyed buildings are estimated in the Kathmandu valley with additional fatalities and damage in villages to N and E

Named location	Modern equivalent	MM intensity	E. decimal longitude	N. decimal latitude	Houses destroyed	Notes
Baghmati	Bungamati	9	85.30	27.62	80	30% destroyed
Bali	?	9			3	
Baneppa	Banepá	9	85.53	27.62	20	
Bara gaon	Baregau	9	85.36	27.6	35	
Bareh	Bore	9	85.38	27.68	20	
Bhat Gaon	Baktipur	10	85.43	27.65	2000	75% buildings (max 4700 destroyed and >6 temples)
British Residency	Kathmandu	8	85.33	27.72	1	
Burba Nil Kanth	Burhanikanth	8	85.36	27.77	2	
Changu Narayan	Cangunarayan	8	85.87	27.69	2	
Changu	Icangu	9	85.27	27.72	20	Old temple destroyed
Chapaly	Capali	9	85.36	27.76	7	
Chappa gaon	Capagau	9	85.32	27.6	35	
Chisapani	Chisapani	8	85.17	27.53	0	Damage to walls of fort
Chitlong	Chitlong	9	85.18	27.67	14	
Deo Patan	Deupatan	8	85.32	27.75	30	Paspatnath temple undamaged
Dharmpur	Dharmapur	9	85.34	27.74	20	
Dharmtuli	Dharamtali	8	85.3	27.76	2	
Dhulaka	Dolacka	8	86.1	27.7		
Digarchi	Xigatse	3	88.9	29.29	0	Not felt
Doti	Doti	5	80.9	29.3	0	
Dukele	Dhulikhel	9	85.58	27.62	21	
Buny Byass hills	?Jumbesi	9			40	
Gorkka	Gorkha	8	84.60	28	2	
Gou Karan	Gokarna	9	85.73	27.73	8	
Hal Chok	Dahacok?	9	85.27	27.72	3	
Handi gaon	Harrigau	9			20	
Harra Sidhi	Harisiddhi	9	85.35	27.62	20	
Hukin gaon	?	9			1	
Kan Sing Choke	?Choksum	9	86	28.07		5 days of foreshocks
Kassa	Dram, Khasa	9	85.95	28		
Kathmandu	Kathmandu	9	85.33	27.72	400	Extensive damage
Kirtipur on ridge	Kirtipur	8	85.24	27.67	14	532 old and frail houses
Kirung	Kyirong	9	85.29	28.48	60	15% destroyed
Kuknah	Khokna	9	85.30	27.63	130	
Kuti (Tibet)	Nyalam	9–10	86.08	28.15	550	83% destroyed
Listi gaon	Listikot	8	85.94	27.89	0	
Lubu	Lubhu	9	85.38	27.62	25	
Mukwanpur	Mukwanpur	8	85.15	27.46	0	Fort damaged less than Chisapani
Mundun Pahar	?E to Tibet	>8				
Nag Desa	Nakdes	9	85.38	27.68	20	
Nala gaon	Nala	9			11	
Narod Devi	?Nardevi	8			1	
Nayakoth	Nuwakot	8	85.23	27.8	3	
Pagah	Panga	9	85.27	27.65	24	
Pahon	Bahongau	8			3	
Palpa	Palpa	6	83.60	27.7	0	
Panouti	Panauti	9	85.53	27.57	19	
Patan	Patan	10	85.32	27.65	285	
Peang	Pyanggau	9	85.33	27.58	8	
Phulam Chock	?Palchok	9	85.53	27.52	300	Temple destroyed
Phurphing	Pharphing	9	85.27	27.61	8	
Sana	?Sanagau	9	85.37	27.63	7	
Sana gaon	Sanagau	9	85.37	27.63	40	
Sangu	Sanga	9	85.48	27.63	8	
Sankhu	Sakhu	9	85.47	27.72	45	Mahadei Temple nearly ruined
Sassanelly	?Sisneri	8	85.40	27.62	2	
Sassanelly Hills	?Sisneri	8	85.40	27.58	20	

Continued

Table 1 continued

Named location	Modern equivalent	MM intensity	E. decimal longitude	N. decimal latitude	Houses destroyed	Notes
Selli gaon	?	9			16	
Shipa	Sipatar	9	85.65	27.75	3	Small houses destroyed
Swambunath	Swayanbhunath	9	85.28	27.71	3	Temple destroyed
Taibu	Thaiba	9	85.35	27.61	18	
Temi	Thimi	10	85.38	27.67	150	
Teshu gaon	?	9			25	
Tewanpur	?Makwanpur	9			10	
Thankote	Thankot	9	85.21	27.68	23	
Tingri	Old Tingri	4	86.5	28.5	0	
Toka gaon	Tocka	9	85.33	27.76	15	
TOTAL number of buildings reported destroyed (includes damage in southern Tibet)					4652	

more than one intensity unit. The 1934 isoseismals shown in Figure 4 are somewhat uncertain within Nepal since they are formed by merging the isoseismals shown in Dunn *et al.*[19] with the intensity data discussed by Pandey and Molnar[27]. Although the resulting extension of the intensity > VIII data northward can be justified using the authority of Rana[28], the position and dimensions of the intensity IX region are conjectural. The centroid of intensity VIII area for the 1833 earthquake is apparently displaced approximately 1° to the west of the equivalent region for the 1934 earthquake, consistent with Khattri's 1987 location. This observation is sustained, however, by three observations only: at Goruckpur, Gorkha and at Chapra (Chuprah) on the Ganges. Examination of the first two of these reports shows that intensity VIII data are but weakly supported by the reports. Thus, were the buildings damaged of poor construction, a lower intensity could be assigned to these villages. The 1833 Chapra account describes slumping ('a chasm of considerable depth formed'), which may indicate a localized region of high-intensity shaking near the Ganges similar to the narrow intensity IX region in 1934. The inclusion of these westernmost VIII intensities in an elliptical fit to this isoseismal results in an 1833 intensity VIII area similar to the inferred 1934 intensity VIII isoseismal. However, a much smaller region for 1833 intensity VIII is admitted by the data if the same pattern of localized severe shaking occurred as in 1934. For example, the sparcity of intensity VIII data in 1833 between Dhankuta and Darbhanga admits the possibility that the area of intensity VIII shaking could be as low as 60,000 km^2 or as high as 100,000 km^2, the latter being the inferred area of intensity VIII shaking in 1934.

Three regions of severe damage were reported in the 1833 and 1934 earthquakes, an observation that led Dunn *et al.*[19] to note a similarity between the two events: the Kathmandu Valley, a high-intensity region near the Ganges, including Monghyr, and a region north of Muzaffarpur. Data for region near the slump belt of 1934, a region of catastrophic lateral spreading, are sparse and although high intensities are recorded near Tirhoot and Purnea (VIII–IX) no liquefaction features are mentioned here in 1833. The area consists of flat-lying sediments and ox-bow lakes across which numerous rivers meander. Further south, at Chapra and Monghyr, accounts indicate localized liquefaction near the Ganges, and minor ground damage is reported to the north, near Bhagmati in the Kathmandu Valley. Shaking intensities in the Kathmandu valley were similar for both the earthquakes, with highest intensities near Patan. Baird-Smith[13] noted that damaging shaking at Monghyr frequently accompanies large earthquakes, an observation repeated by both Dunn *et al.*[19] and Pandey and Molnar[27]. A favoured explanation for localized high intensities in this region is that surface waves are amplified in the water-saturated sediments as they approach the southward shelving bedrock surface south of the Gangetic Plain[27]. Similarly, the lake deposits of the Kathmandu valley can be assumed to be responsible for localized high intensities and rapid variations in intensity in this region as were observed most recently in 1988 (ref. 29).

Magnitude of the 1833 earthquake

Various empirical relations have been developed to relate felt areas to magnitude[30–32]. A physical formulation to account for the observed relation between intensity and moment magnitude has been proposed by Frankel[33] the coefficients of which have been fit to the global database by Johnston[34]. I shall refer to this as the F94 model. Moment magnitude, *Mo*, of an earthquake in the F94 model is related to the area enclosed within a specified isoseismal intensity contour, *S*, with an expression of the form

$$\log Mo = a + b \log S + c\sqrt{S}, \quad (1)$$

Table 2. Summary of the 1833 intensity data from Indian newspapers (see Appendix) compared to the 1934 Bihar earthquake intensity data from Dunn *et al.*[19]

Location	Mercalli intensity 1833	Mercalli intensity 1934	Longitude (dec. E°)	Latitude (dec. N°)	1833 Shaking duration (s)	1833 Shaking azimuth
Main shock						
Agra	V	VI	78.02	27.13	5	
Allahabad	VI	VII	81.92	25.43	120	
Arrah	VIII	VII	84.67	25.57	60	N
Bankipur	VIII	IX	85.1	25.6		
Bankura	VI	VI	87.07	23.23		
Benares (Varanasi)	VII	VII	83.02	25.3	150	NW
Bhagalpore	VII	VIII	87	25.25	480	
Buxar	VII	VII	83.98	25.57	180	N
Calcutta	V	VI	88.4	22.57	120	N
Chandernagore	VI	V	88.4	22.73	180	
Chapra	VIII	IX	84.73	25.78		N
Comercolly (Kumarkhali)	V	VI	89.18	23.87		
Dinapore	VIII	VIII	85.05	25.63		
Delhi	IV–V	V	77.2	28.7	300	NE
Futtigurh	IV	VI	79.6	27.3		
Gazeepore (Ghazipur)	VI	VII	83.5	25.6		
Goruckpore	VIII		83.4	26.8	180	E
Hathras	V	VI	78.05	27.6	60	NW
Hazaribagh (Hazareebad)	IV	VI	84.27	23.95		
Jabalpur	II–IV	V	75.98	23.33	(ref. 12)	
Kathmandu	VIII–X	IX–X	85.33	27.68	60	NE
Kirong (Kyirong)	VIII		84.25	28.45	(ref. 2)	
Lhasa	I		91.17	29.68	(ref. 2)	
Lucknow	VII	VI	80.93	26.87	180	
Lohooghat (Lohughar)*	II		80.08	29.38	(ref. 13)	
Meerut	V	V	77.72	29.02	15	
Monghyr	IX	X	86.56	25.38	60	NE
Mullye (Murliganj?)	V		80.72	26.92	350	N
Murshidabad	V	VI	88.4	24.2		
North Tirhoot	VIII	IX	85.4	26.4		
Nusserabad (Nasirabad)	I	IV	90	24.43		
Patna	VIII	IX	85.1	25.6		
Rhotas Hills	IV	VI	83.83	24.63	120	
Purnea	IX	IX	87.47	25.77	NE	
Rungpore (Rangpur)	VI	VI	89.3	25.75	40	
Tingri, Tibet	IV	V	86.5	28.5		
Tirhoot (N. Tirhoot)	IX	IX	85.4	26.23	60	E
Tirhoot (Muzaffarpur)	IX	IX	85.4	26.12	60	N
Xigatse, Tibet	I	V	88.7	29.3	(ref. 2)	
Aftershocks						
Berhampore 4 Oct	V		86.1	300	0	
Jaunpur 9 Oct	V		82.68	25.75	3	
Malda 4 Oct	V		88.13	25.03		
Goruckpore 18 Oct	VII		83.4	26.8	60	E
Bhagalpore 4 Oct	VII		87	25.25	10	
Purneah 4 Oct	V		87.47	25.77		

*Baird-Smith[13] citing McCleland, p. 1051 (note that due to misnumbering, a second p. 1051 occurs, and p. 1053–1056 appear thrice distinguished by * and † symbols).

where the constants a, b and c are determined empirically for each isoseismal area. Intensity magnitudes are shown in Table 3 using the coefficients derived from worldwide data that include six Indian events[34]. A mean magnitude determined for the 1833 earthquake in this way is 7.5 ± 0.3, close to the $M = 7.6$ adopted by Khattri and Tyagi[1]; however, a monotonic decrease in estimated magnitude with intensity is observed, presumably indicating systematic differences between the global coefficients and those appropriate for the 1833 event. Data for several Indian earthquakes are presented in Figure 5 to illustrate these possible biases. Notwithstanding the ensuing discussion, a cursory inspection of Figure 5 indicates that the 1833 earthquake must be greater than

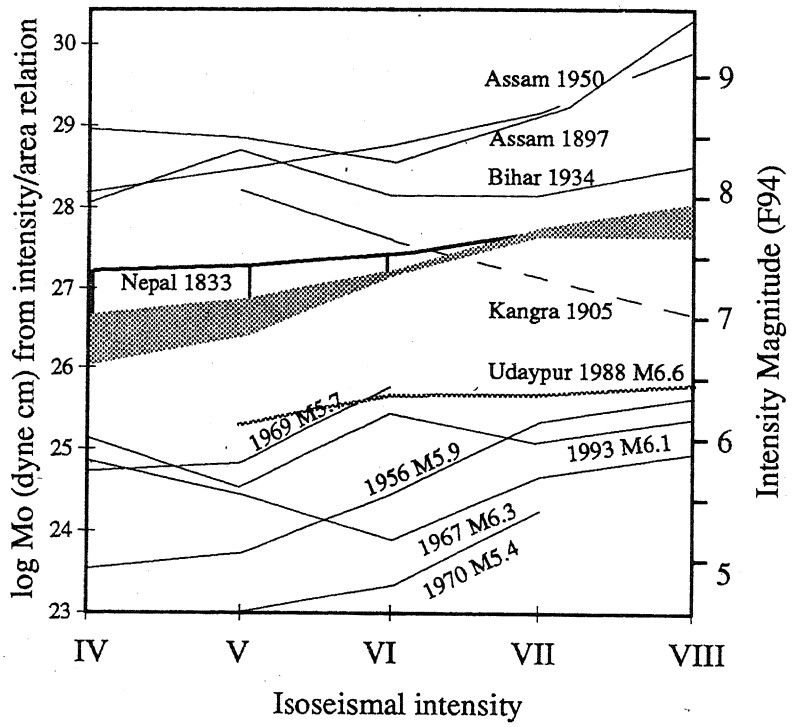


Figure 5. Moment magnitudes inferred from Johnston's F94 coefficients for several earthquakes in Nepal and India. The shaded area represents the range of magnitudes inferred for the 1833 earthquake using elliptical isoseismals (Figure 4). The overlying fenced region assumes circular isoseismals (see text). Assam 1897 data from Oldham[18], Bihar 1934 from Dunn *et al.*[19] and Pandey and Molnar[27], and Udaypur from 1988 data from Dikshit and Koirala[29] and Sinha[26]. Events in tilted script from Johnston[35].

Table 3. Isoseismal areas and estimated movement magnitudes for the 1833 earthquake. The constants *a*, *b* and *c* are from Johnston[35] for global data least-squares fit to equation (1). Moment magnitudes for elliptical areas (F94E) are from Figure 4, and those for circular areas (F94C) for diameters equal to the long axes of the ellipses in Figure 4 were unconstrained by data to the SE. Areas are expressed as $\log_{10}$(area km^2) and *Mo* in $\log_{10}$(dyne cm).

	Intensity felt	IV	V	VI	VII	VIII	Mean *Mo*
a	19.62	18.36	20.44	19.51	22.96	24.3	
b	0.5	0.903	0.607	1.307	0.628	0.352	
c	0.00163	0.00206	0.00312	0.00079	0.00307	0.00655	
Elliptical area (E)		6.16 ± .08	5.87 ± .06	5.55 ± .02	5.35 ± .02	5.02 ± .04	
log(*Mo*) F94E		26.4	26.7	27.2	27.8	28.0	
Mo F94E		6.9 ± 0.2	7.1 ± 0.1	7.46 ± .02	7.82 ± .03	7.9 ±0.3	7.5 ±.3
Circular area (C)		6.42	6.16	5.86			
Mo F94C		7.6	7.9	7.9			7.8 ± .2

$M = 6.6$ and less than $M = 8.1$, the inferred magnitudes of the Bihar 1934 and the Udaypur 1988 earthquakes.

Frankel[33] assumes that attenuation is radially symmetric, resulting in circular isoseismals. These are not appropriate for the Himalayan plate boundary. Yet it is not clear whether the observed isoseismal ellipticity is an artefact of reporting, an east–west amplification effect, or a north–south attenuation effect. The elliptical isoseismals inferred in Figure 4 have an approximate 2:1 aspect ratio with their long axes parallel to the Himalaya, but, as mentioned previously, this is possibly biased by the absence of reports to the north and to the SE. By assuming that the lower-intensity isoseismals approximate circles with their radii equal to the semi-major axis of the ellipses shown in Figure 4, circular intensity IV, V and VI isoseismals can be invoked that yield intensity magnitudes of 7.6, 7.9 and 7.9, respectively, raising the mean magnitude intensity to $M = 7.8$. Although the observations permit this interpretation towards the SE, the 'not felt' data from Lhasa and Xigatse[2,12] are inconsistent with circular isoseismals. Moreover, the elliptical aspect of Himalayan isoseismals

appears to be a common feature of Himalayan earthquakes. Intensity VI–VIII isoseismals for the $M = 6.6$ Udaypur 20 August 1988 earthquake exhibit an aspect ratio of approximately 4:1 (refs. 29, 36). Isoseismals VI–VIII for the 1991 Uttarkashi earthquake are elongated 2:1 along-arc[37] and intensity > VI and VII isoseismals for other Himalayan events show elongation parallel to the arc (e.g. Kangra VII 1905 and Bihar 1934). Thus, although the absence of radial symmetry for intensities > VIII could be ascribed to source geometry, it is surely necessary to invoke transmission effects as the reason for the elliptical isoseismal areas VI–VII found in the southern Himalaya and Gangetic Plain. For example, it is likely that the sediments overlying the stable continental crust in the Gangetic Plain act as a waveguide for surface waves travelling east–west with periods of 1–3 s (ref. 38). Similarly, the dipping interface between the Indian plate and the overlying sediments may result in less interference along strike than updip and downdip, favouring along-arc perceptibility. For the lower intensities that are felt south of the Gangetic Plain the intensity distributions for the four great earthquakes between 1897 and 1950 are approximately circular, consistent with transmission within a uniform crust.

Epicentre, aftershocks and rupture areas in 1833 and 1934

The location of the main shock cannot be determined unequivocally from the data presented here. This is largely because most of eastern Nepal and southern Tibet are unrepresented in the felt report data, preventing incontestable closure of isoseismals greater than VIII in these directions. A tantalizing intelligence attributed to 'vulgar report' by Campbell[12] is that for 5 days before the earthquake 'noises similar to the firing of cannons were heard as if underground: and in this neighbourhood' (Kan Sing Choke, possibly Choksum at 86°E, 28.07°N), 'the high road to Lhasa is said in many places to be blocked up by the fallen earth from the mountains'. If these are interpreted as foreshocks close to the epicentre of the main shock, the epicentre would be placed unreasonably close to a rapid decay in felt intensities eastward. That is, intensities at Tingri 100 km to the NE are estimated at IV, whereas at Kyirong and Kathmandu, similar distances from Choksum to the NW and SW, intensities exceed intensity VIII. Thus, if the main shock, or one of the foreshocks, occurred near Choksum, this may have been the eastward limit of subsurface rupture. The few aftershocks that were reported in 1833 do not provide any clear idea of the possible rupture zone of the earthquake southward or westward although their widespread manifestation favours a shallow thrust more than a deep focus event such as the 1988 Udaypur earthquake. Continuous ground motion in the days following the 1833 earthquake was claimed at Monghyr, Tirhoot and Kathmandu, the three regions of highest-intensity shaking. The concentration of felt aftershocks within these areas may not indicate the proximity of aftershocks to these locations, but instead may indicate either that the frequency of shaking in these locations was optimal for people to sense (e.g. 2–4 Hz, ref. 33) or that anxiety about further destruction heightened their awareness to aftershocks. On 4 October 1833, a strong aftershock was felt simultaneously at Kathmandu, Allahadad, Berhampore, Malda, Purneah and Bhagalpore, and another on 18 October at Goruckpore, Kathmandu and Allahabad. A surprising number of aftershocks were felt in the slump belt region in 1934 and the largest instrumentally located aftershock in 1934 is also found in this region[19].

Seeber and Armbruster[22] and Chander[39] propose that the 1934 earthquake ruptured a region forward of the main frontal thrusts of the Himalaya, at least as far south as the slump belt. High-intensity shaking was also observed near the Greater Himalaya and central Nepal close to the instrumentally determined epicentre, approximately 100 km north of the slump belt[27]. If it is assumed that the 1934 earthquake was $Mo = 8.1$, its rupture area could easily extend and embrace these two regions and extend an equivalent distance, or greater, along strike. However, the 1833 earthquake with a moment magnitude of 7.6–7.8 is unlikely to have ruptured an area more than 70 km on a side. The absence of catastrophic damage at the forts of Chisapani and Mukwanpur argues against a rupture zone centred at or near latitude 27°N, and it seems unlikely that rupture further south (beneath the plains) could generate >30% destruction in towns in southern Tibet. Thus, the rupture zone of the 1833 earthquake appears to have been centred in northern Nepal or southern Tibet, and unless it was a thrust of unusual aspect ratio, it would not extend beneath northern India. Thus, the inferred intensity VIII and IX shaking in the Gangetic Plain may have resulted from basin resonance, and not from the underthrusting invoked to have occurred during the 1934 event.

Following this line of reasoning, if a $M = 7.7$ 1833 earthquake near the Greater Himalaya could generate intense localized shaking in the Gangetic Plain, so too could the 1934 Bihar earthquake. Could the slump belt in 1934 be a product of basin resonance and prolonged shaking, thus not requiring slip to extend beneath the Gangetic Plain? Catastrophic lateral spreading occurred in the 1964 Alaska earthquake and was the result of vigorous and prolonged shaking of saturated sediments. It is not difficult to imagine that catastrophic lateral spreading could be confined to a region of especially intense shaking in 1934. The removal of the requirement for slip terminating near the slump belt would remove the obvious difficulty that there is no evidence for a

topographic bulge above the required southern edge of the inferred rupture. This was not revealed by levelling, nor is there any topographic evidence to suggest that the ruptures of previous shallow thrust earthquakes have terminated near the slump belt.

If it is assumed that the 1833 and 1934 earthquakes were both shallow thrusts, a limited number of geometries are possible if we require additionally that the two ruptures do not overlap. However, this requirement is not essential because the rate of development of available slip from Indian/Tibetan plate convergence is 1.5–2 m per century[40], and slip released in the 1833 event could have been effectively renewed by 1934. The simplest relation for the two events is that a region north of Kathmandu ruptured in 1833 and that a region roughly four times the area ruptured in 1934 immediately to the east. However, alternative and not necessarily more complex rupture geometries can be envisaged that would also be consistent with the available data.

Isoseismal areas and great Himalayan earthquakes

Were the worldwide coefficients determined by Johnston[35] for equation (1) appropriate to relate the felt areas of Indian earthquakes to their moment magnitudes, each event in Figure 5 would be represented by an horizontal line. However, intensity magnitudes determined for the four largest events show significant variance, indicating that the data for these events are unreliable, or that the Eurasian/Indian plate boundary requires adjusted coefficients in equation (1). Unfortunately, the intensity data for the four great earthquakes are both confusing and incomplete. They are confused partly because intensity scales developed for the industrial nations are not easily applied to the type of construction and destruction in each area, and partly because those responsible for the compilation of intensity data were unable to visit much of the afflicted areas to examine inconsistencies in written reports and hence apply uniform criteria to the observed damage. They are incomplete because isoseismal coverage is typically only for one-half of the shaken region. The data in Figure 5 have area uncertainties of up to 50%.

For all the four great earthquakes the maximum 'felt area' isoseismal reported in the standard works on these earthquakes[18, 19, 41, 42] yields moment magnitudes 1–2 magnitude units lower than the magnitude derived from the intensity IV isoseismal. Given that the intensity IV–VI isoseismals yield moment magnitudes consistent with those derived from radiated seismic energy, a possible conclusion is that the 'felt area' isoseismal has been significantly underestimated for these events. Surface wave dispersion results in progressively lower frequencies dominating the accelerations at increasing distances from the epicentre and distant accelerations are presumably shifted to frequencies that are not optimum for human perception. Distant 'felt' reports thus become increasingly 'instrument'-based: long-period oscillations in reservoirs, movements of chandeliers in churches, and slow movements of doors.

The 1897 isoseismals documented by Oldham[18] consist of precisely delineated regions of damage to masonry structures, and approximate areas assigned grouped Rossi–Forrel intensities. The low-intensity isoseismal areas are completed by extending Oldham's elliptic isoseismals to the north; however, the resulting areas may be uncertain by 50%. Isoseismal areas listed by Tandon[42], Poddar[43] and Ray[44] for the 1950 Assam earthquake yield systematically low moment magnitudes unless these are doubled to include the areas not visited to the north and east of the epicentre. This extrapolation is undertaken 'with the risk of great departure from the truth'[44], a truth already uncertain for the same reasons that the 1897 isoseismals are unreliable. Estimates of lower and upper bounds for these isoseimals result in more than a ±0.5 variation in assessed moment magnitude. Both the Assam earthquakes yield mean moment magnitudes between 8 and 9.

The 1934 Mercalli intensity data reported by Dunn *et al.*[19] yield reasonably consistent moment magnitude estimates although the values are somewhat higher than those derived by Johnston[35]. This is presumably because of uncertainties in closing intensity IV–VI isoseismals in southern Tibet. The intensity VIII area adjusted in Figure 4 to merge the felt data from north and central Nepal[27] with the data from Dunn *et al.*[19] yields a good approximation to the observed moment magnitude.

The Kangra isoseismal data as reported by Middlemiss[41] are unique among the four great Himalayan events in that they cannot easily be reconciled with the Johnston F94 coefficients. In particular, the mapped intensity VIII data are confined to two small regions, yielding an absurdly low estimate for moment magnitude ($M \approx 7$) compared with hitherto accepted magnitudes for the event ($M = 8.4$). This anomalously small area is evident in Figure 1. Yet the estimated intensity V isoseismal area for this event is consistent with an $M \geq 8$ earthquake. Even if the two mapped intensity VIII areas were to connect, which Middlemiss[41] is adamant they do not, or the Kangra intensity VIII isoseismal to bend southward, as suggested by Molnar[45], the increased area would bring the moment magnitude to no more than $M = 7.5$. A localized intensity VIII area associated with a large earthquake would normally be interpreted as being symptomatic of a shallow earthquake, or an earthquake in a region with abnormally high attenuation. Yet no surface rupture was reported there, and the widespread felt area indicates that the magnitude of the earthquake is typical of other great Himalayan events. The inversion of levelling data from the easternmost area is consistent with a slip on a shallow thrust of 7.5 m

(ref. 46), again consistent with the occurrence of a great earthquake. A possible explanation for a small high-intensity region and a large felt area is that the Kangra 1905 event was a slow earthquake. Slip on a shallow dipping buried thrust may have occurred at rates lower than the 1897, 1934 and 1950 earthquakes, insufficient to generate a widespread region of high accelerations but sufficient to stimulate long-period waves, resulting in a large felt area. In this interpretation, the high intensities observed near Kangra and Dehra Dun may have been associated with secondary shallow faulting in response to the underlying slip.

Conclusions

A 1 million km^2 of northern India was shaken by a significant earthquake in 1833 that appears to have been centred in north central Nepal. Although intensity X shaking occurred locally and more than 4000 buildings are reported destroyed, the loss of life (≈500) was small largely due to the occurrence of two large foreshocks hours and minutes, respectively, prior to the main shock. A possible location for the epicentre of the 1833 main shock is approximately 50 km north, or northeast of Kathmandu, although the limited number of observations in east central Nepal and southern Tibet permit an epicentre to the east of Kahthamandu, close to the epicentre of the 1934 Bihar earthquake. The 1833 earthquake resulted in damage intermediate in severity between the 1988 Udaypur $M = 6.6$ and 1934 Bihar $M = 8.1$ earthquakes. A mean moment magnitude of $Mo = 7.7 \pm 0.2$ is obtained by applying Johnston's F94 relation between isoseismal area and moment magnitude.

The inferred $M < 8$ magnitude and consequent relatively small rupture area of the 1833 earthquake indicates that it contributed insignificantly to reducing the potential slip in the Central Himalayan Seismic Gap between the Kangra 1905 and Bihar 1934 rupture zones. The slip associated with the event may have been 1–2 m, an amount consistent with a renewal time of 100 years. This and its proximity to the inferred rupture zone of the 1934 Bihar earthquake suggests that if the 1833 event occurred on a thrust fault, it may have ruptured a region abutting or overlapping the 1934 rupture. However, although remarkable similarities between the 1934 and 1833 isoseismal areas exist, the mechanism of the 1833 event is unknown.

The application of the F94 Johnston[35] algorithm to intermediate-intensity isoseismal areas of the four great earthquakes in the Himalaya yield moment magnitudes consistent with the currently accepted values. Maps for isoseismal intensities less than V and greater than VII are either less reliable or incomplete for each of these earthquakes. Even when generous allowance is made for the northward extension of the perceptible felt area, the intensity II area for the four events yields values that are too low by more than 1 magnitude unit, indicating that recalibration of the F94 algorithm for Indian earthquakes is desirable.

The 1905 Kangra earthquake is unique in that its intensity VIII felt area is anomalously small for a great earthquake. The seismically determined magnitude, the geodetically estimated slip, and the intensity IV–VI felt areas are consistent with an $8.5 > Mo > 8$ event, whereas the intensity VIII area is appropriate for an $Mo \approx 7$ event. A possible explanation for this anomalously small area of intensity >VIII shaking is that the earthquake occurred with a substantial slow component, and that the observed high intensities were associated with secondary ruptures near Kangra and Dehra Dun. The absence of broad-band seismic data for this and other Himalayan earthquakes makes such a conclusion tentative. However, if slow earthquakes do occur in the Himalaya they would have the benefit of absorbing plate convergence with less high-intensity shaking than normal high-frequency earthquakes. Moreover, if other $M \approx 7$ earthquakes in the historic record represent the secondary seismic manifestations of slow great earthquakes, the maturity of the Central Seismic Gap may not be as advanced as is currently believed.

APPENDIX

Accounts of 26 August 1833 Nepal earthquakes and aftershocks

Times in the original articles are converted to a 24-hour clock. Note that these original times are generally unreliable because they are expressed in local time usually set by a local sun clock and span a one-hour time zone. A precise astronomical clock was stopped in Calcutta by the second foreshock at 23:34′48″. The main shock occurred at approximately 23:54 and the first foreshock at 17:00. Mercalli Intensities are assigned to the main shock and to the five reports of aftershocks. The first line of each entry indicates the estimated Mercalli Intensity, where this can be determined.

Three reports from Nepal containing the observations of Campbell[2,12] are included in addition to an anonymous commentary from Agra.

Agra V

The Earthquake – Several shocks of an earthquake were felt in cantonments, as also in the city of Agra, on the night between the 26th and 27th instant. The shocks came in rapid succession, were strong, and lasted a few seconds each, but we have not heard that any accidents have happened.

(*The India Gazette, Calcutta, Tue., 12 Sept. 1833, III, 872, reprinted from the Mofussal Ukhbar, Aug. 31, 1833*)

Agra. The Earthquake.

The following account from a correspondent in Agra summarizes accounts in the above newspaper reports south of latitude 26°N, and adds a few locations not encountered in the primary articles that follow. It is included because it appears to be a systematic list of 1833 settlements from which news of the event was expected. It

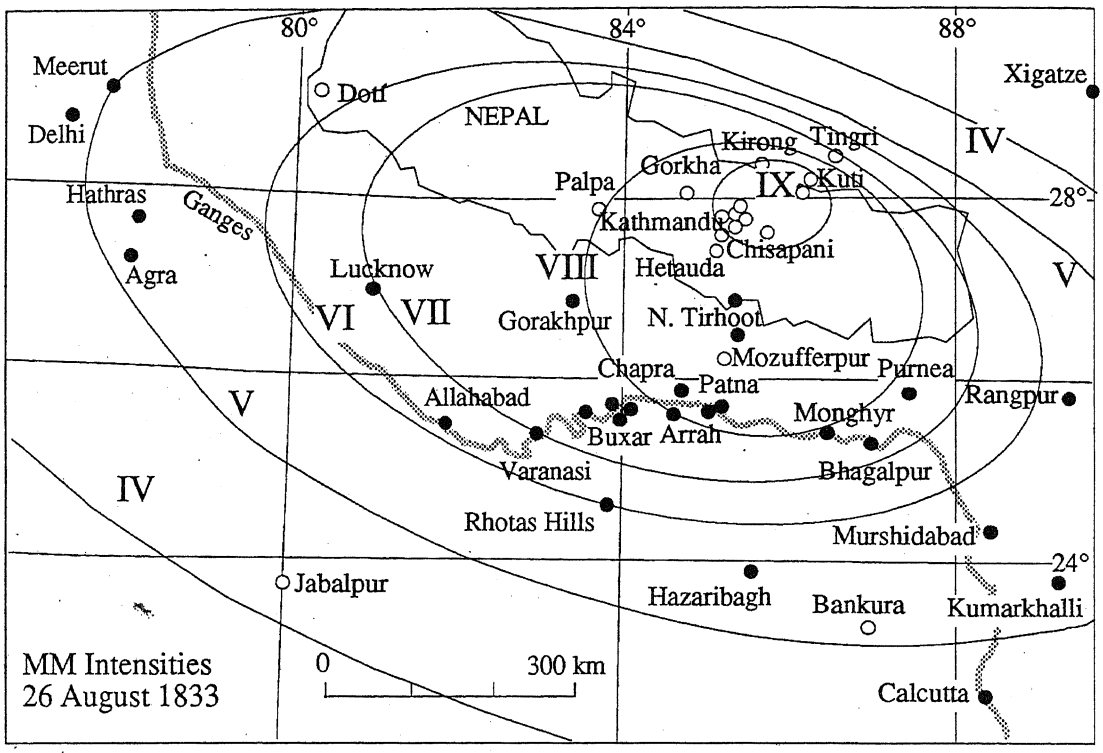


Figure 6. Locations reporting the 1833 earthquake. Solid circles from Indian newspapers and journal articles. Open circles from Campbell[2,12]. Location coordinates are listed in Tables 1 and 2.

claims to be the first of the two reports but a search for the second, which purports to describe damage north of 26°N has been unsuccessful. Mullye is not mentioned in this report, suggesting that it is located north of latitude 26°N.

We are happy to perceive a hint in the Calcutta Courier that the Secretary to the Asiatic Society is likely to favour the public with an account of the earthquake of August 26. To assist in rendering this as complete as possible, we have been induced at this early period to take a review of the information already collected for the purpose of showing what is yet wanted from individuals residing at places where the shock was sensibly felt. We would strongly recommend that a set of questions be circulated through the government by the Asiatic Society, as it is probable that there are many subjects connected with the sciences, besides the phenomena attending the earthquake, on which a light may be thrown by means of such an enquiry. We intend in the first place to confine our remarks to the sensible effects of the earthquake in different situations, and here we shall be obliged to recapitulate facts with which most of the readers are already familiar. The earthquake must have proceeded from a point or from an extended line or a space. The first subject, therefore, seems to be determination of these limits; in order to do this, we ought to possess properly an understood measure of intensity, so that we might ascertain our vicinity to the vortex or centre. Supposing that we assumed the earth to be of equal solidity throughout its surface, and that the earthquake had proceeded from a certain point, we should find a regular succession of accounts which would be exactly similar at similar distances. If the point, therefore, is ascertained, and any discrepancy appears in the accounts, we are obliged to refer this to the quality of the intervening strata. We may premise that there were three severe shocks; one at sunset, another at a little past 11, more severe, and the third one at or near midnight, the severest of all. Now did these three happen at the same place or at several places? If they happened at the same place, we should, by commencing with the accounts from the eastward remote from the epicentre, find exactly similar accounts to what we would have from the westward, and the difference of strata would have no avail, for having obtained three different degrees of intensity, they would all shade into obscurity (if we may be allowed to use the expression) in the same order at proportional distances.

22°N: We shall proceed from the southernmost point from which information has been obtained and take parallels of longitude from east to west. From the westward [this should be eastward?] of Calcutta no recounts have been received. At Calcutta and Dum Dum 2 shocks were felt, the first comparatively light one at 23:36, which stopped the astronomical and other clocks. The second one came on at 23:56, occasioning the doors to rattle while the punkhas, the wall shades, etc., were all thrown into violent agitation. It lasted about 1 min. This is all that has been recorded for this parallel, no accounts having been received from Midnapore, Singboom, Oodeypore, etc.

24°N: In 24°N latitude there are no accounts from Dacca – At Comercolly two shocks were felt, the first one at 23:45 lasting nearly a minute, the second one, somewhat lighter, about 10 min after. At Moorshedabad (=Berhampore), the first shock was felt at 19:00 but two others about midnight within 10 min of each other shook the house, awakened the birds, and caused the hanging lams to vibrate for 4 min afterwards. So severe was the shock that the writer feared his house would fall. The earth trembled so violently that it was with difficulty one could keep his footing.

At this interesting point, our information terminates, and we pass on to Hazareebad (Hazaribagh) on the southern side of the concussion. At this station the early shock at 19:00 was not felt; but one at 23:30 and a second one half an hour afterwards lasted 2 min, causing the punkha to swing and the door to rattle loudly. In a letter from the Rotas Hills, the first shock was felt at 23:10, a second rocking came soon afterwards, and at 23:11 it returned more severely, lasting not

less than 2–3 min; the roof of the house creaked, and the doors and windows rattled.

There are some discrepancies in this parallel which might perhaps have been explained had we possessed more extended accounts. For instance, the slightness of the last shock at Commercolly; secondly, the sudden increase at Moorshedabad. The slightness of the effects at Hazareebag, and the severity, considering the remoteness under the Rotas Hills we should presume that the situation of the last account must have been made considerably to the northward of Hazareebag.

25°N: We shall now take the parallel 25°N from whence the fullest statements have been obtained. From Sylkat we have no accounts, our first are from Rangpore, nearly in 25°N, where the shock was felt as at Moorshedabad. The first one took place at a little after sunset, slight and of short duration. The second one at 23:30 was strong, lasting about 30 s and the third one was still more severe, of 40 s duration.

No accounts have been received from Dinagapore, Malda or Raymahil (Rajmahal), and we are obliged to pass on to Poornea, where the shock produced still more severe effects. The first shock was felt at 17:00 slightly, the second one at 23:00 was severe and the third one at 23:18 was the most violent. The birds were driven from their nests; men were scarcely able to stand; cattle ran wild with fright; several old buildings in the town tottered to pieces; a part of one in the station sank partially; a heavy cornice fell, every building suffered more or less, and a recently vacated upper-roomed house had not an arch left. Large masses of bricks actually fell out of the walls. The roofs in many places split, and the beams and Burgahs separated an inch and more. The out-offices in the compound became a heap of ruins. The shocks are mentioned as occurring from 23:30 to 08:00. The last was somewhat severe; no mention is made of a subterranean sound.

Our next accounts are from Bhagalpore. The first shock was felt at 23:15, and a tremulous motion was felt for 3–4 min; at about 23:45 the third shock was felt when the chairs moved and the door rocked about; this lasted for about 8 min. Between this and daylight three other shocks were felt, but none equal to the first one; a rumbling noise was also heard, but not very loud.

From this we proceed to Monghyr. At this station the first shock occurred at 5 min after sunset accompanied by a rushing noise shaking like the movement of a steam engine. The second took place at 23:25, the third one at midnight, the fourth one at 02:00, and the fifth one at 03:45, etc. The second and third ones were most severe, causing considerable damage to houses, in which 34 were more or less injured by the falling of the balconies and balustrades, and by the cracking and separation of the walls and beams. The ground seemed to rise in waves about a foot in height. The lamps were all extinguished, making with the glasses on the sideboard, agitation of the jilmils and rumbling underground which accompanied the earthquake made a frightful noise. The first shock seemed to last for 30 s, and the second and third ones about 1 min each. Mr. H. and Colonel B., Mr. Pages and Captain Watson's houses sustained considerable damage. Large portions of the Rampart of the fort fell down. The shock was distinctly felt by the people in boats. Another slight shock occurred at 21:00 of the 27th. The Commissioner's house, situated at the highest point of the station, suffered the least, though the earthquake was severely felt there.

From Monghyr we proceed to Patna having received no accounts from Bar. The first shock was felt there at about 18:00, rather slight. The next one was felt when a number of houses fell and various sorts of injuries were sustained. Rajah Khan Bahadoor's stable fell, by which 7 of his horses were killed. Capt Elliot's house was also levelled with the ground. E. DeCosta's house cracked. At 23:30 says one, I was awakened by a horrid rattling and shocks, apparently from the east to the west with a rumbling noise, the servants running into room in great consternation. The last shock occurred at midnight, which was the most severe. At Bankepor (=Bankipore), 6 miles from Patna, houses were cracked, amongst which were the magistrate's Kutchery, and in the city the opium godown shared the same fate as also [the] gateway built by Maha Rajah Mohur Jay Singh, while many pucka dwellings had entirely fallen. Eighteen shocks were felt afterwards but they appeared continuous; the last one was at 08:00. At Dinapore (Danapur), 10 miles away from Patna, the former Brigade Major's house was split.

At Arrah, a little east of Dinapore, the first shock was not felt at 23:15, one occurred lasting a couple of seconds, the second one began about 10 min afterwards and was the most severe of the two. It commenced by a slight tremulous motion which gradually increased to the most terrific undulations, and was accompanied by a load rumbling noise, like thunder, underground. Doors, windows and every other article of furniture were thrown into a most violent agitation, several walls fell and some roofs were thrown off.

At Chuprah, in nearly the same longitude but a little to the north, the first shock at 17:00 was felt but slightly; at 23:00 another one came on accompanied by a tremendous rumbling, and a violent rocking of the house. At 23:15 the severest shock commenced with a tremulous motion, and a noise like thunder, followed immediately by a very long and awful agitation of the earth, which vibrated in a horizontal direction north and south. Slight shocks continued until 08:00 and a tremulous motion at intervals during the day. Large pieces of plaster fell from the walls in the rooms; native dwellings were destroyed; 2 miles from thence the earth opened; in consequence of the injury the court house sustained, the Cutchery was then held in a tent.

At Buxar, the force lessened – several smart shocks were felt, the first decided one took place at 23:20 but the smartest one occurred at 23:55. The motion was so strong as to upset several wall shades and glasses in one of the Bungalows there and the general rattling of the doors and windows and the creaking of rafters were really most alarming. The indigo factory was cracked, a rocking motion was felt by an officer in a budgerow; 5 miles above Buxar at Koorantadhee immediately opposite Buxar, it was so slight as to scarce waken a single individual.

At Gazeepore, the first shock felt was at 23:00 and the second one at an interval of 40 min. The first one was slight, the second as it were 4 or 5 successive undulations accompanied by a subterranean noise, heard however by very few. The cots of the soldiers were violently agitated so as to cause the occupants to start from their sleep, the walls and joints cracking.

At Benares, the first shock at 18:00 was not felt; at 23:00 a shock took place which continued for 4–5 s. This was succeeded by another one not very violent for 4 or 5 s; a more severe one was felt afterwards which lasted 2.5 or 3 min and the earth ceased to shake gradually.

At Allahabad, only two shocks were felt; 23:30 the same phenomenon presented itself with a more striking character, the short but quick vibration continued for nearly 2 min, accompanied with noise from the doors, windows, etc., and a dreadful hollow sound from the river.

We wish we could have completed this with a report from Culinger.

Having now finished an account of our principal parallel, we shall proceed to draw the inferences from the limited facts which have been collected. Rungpur and Moorshedabad appear equidistant from the centre. The former is equally far from Purneah as Moorshedabad is from Monghyr while at the latter two stations, the shock appears to have produced equally severe effects. At Bauglepore, although very near to Monghyr the effects are wonderfully lessened, leading one to the idea that the ground under Bauglepore was in no way connected with the earthquake, that it was more solid, etc., and therefore that the tremor lasted longer. A sound was heard at Bauglepore, Monghyr and less so at Purnea. The rumbling of stones, the rush, as it were of water, and of a mighty wind, are accompaniments of earthquakes which will probably lead eventually to the discovery of their cause. The appearance around Monghyr as well as the hot springs there testify to the volcanic nature of the subjacent strata.

At Patna the effect seems still nearly as severe, apparently diminishing towards Arrah. Our previous accounts from Hazareebag indicate that the effects decrease southward, while those from Purneah on the one side and Chuprah on the other indicate the boundaries east and west showing also severity in the latter case, that the ... effects of the earthquake are to be... in a northern direction from these stations. We observe the gradual diminution of the effects... Arrah until we come to Allahabad and may almost conclude that the last shock was just perceptible at Calinger.

By comparing Moorshedabad and Rungpore, we find them equal, and by comparing in the same manner Allahabad and Calcutta we shall also find that the effects are similarly described and that they are at equal distance, respectively, from Patna and Monghyr. The facts then lead to the conclusion that a given concussion produces similar effects throughout the earth's surface at equal distances. Should such be the case, and our information here be accurate, we should with these data be able to solve the following quotation: *Given the effects of an earthquake at two known points, and also the direction required* (missing words in the printed version of this statement), *the point from whence the concussion arose...* With this idea we have measured the distance between Calinger and Patna, and assuming that at Agra the effects were similarly felt as at Calcutta, we have put one leg of the compasses on the map at Agra and swept a circle which passes through Gazeepore, Guruckpore up the course of the Gandack to its origin in the snowy mountains.

Now we have already had an account of ... so that the effects of the earthquake must according to our data have been more northerly. It was, however, felt at Delhi, in a similar manner, and if we take the same radius we shall still keep to the source on the Gandack, from which we have not as yet obtained any accounts. We shall reserve another parallel for our next number.

(*Reprinted from Mofussul Akbar, 21 Sep. 1833.*In *The Bombay Courier*, **476**, 8 Oct. 1833)

Allahabad, 26 August **VI**

At Allahabad we regret to hear that death has lately been truly alarming and frequent in its occurrence [cholera epidemic]. We are also informed that two smart shocks of an earthquake were felt there during the night of the 26th. Our informant, who resides at the fort, states that he was awoken from a sound sleep at 23:00 in consequence of the tremulous motion being communicated to him by the bedstead – at 23:30 the phenomenon presented itself again with a more striking character; the short but quick vibrations sensibly continued for nearly 2 min accompanied with a noise from doors, windows, etc., and a dreadfully hollow sound from the river.

(*Reprinted from the Calcutta Courier in the Bengal Hurkaru and Chronicle XII(58), New series, Thu., 5 Sept. 1833*)

Allahabad, 16 Sept. **VI**

August 26. First shock 5 min before sunset, not severe; close, hot, light airs from NE; barometer 29.273. At 04:26, barom. 29.213. On the 25th same hour 29.253; thermometer 88°F on both days. About 11 PM and 11:30 very severe shocks were felt; loud noise, resembling distant thunder was heard. The river Jumna was thrown with violence against the banks. Many persons left their houses, both Europeans and natives.

(*The India Gazette, Calcutta, 24 Sept. 1833, II, 882*)

Arrah, 27 Aug. **VII**

Last night the inhabitants of this town were thrown into a state of alarm by two severe shocks of earthquakes. The first one took place at 11:15 and lasted for about 10 s. The second one began in about 10 min and was the most severe of the two. It commenced with a slight tremulous motion which gradually increased to the most tremendous undulations, and was accompanied by a loud rumbling noise, like thunder, underground. It appeared to proceed from north to south. Doors, windows and every article of furniture were thrown into the most violent agitation: several walls fell and some roofs were thrown off. All rushed to the streets and the whole town was in a state of commotion. We understand that some of the *sahib log* abandoning their beds ran into the streets in great consternation.

(*Bengal Hurkaru and Chronicle XII(58), New series, Thu., 5 Sept. 1833, reprinted from the India Gazette Calcutta, Wednesday, 4 Sept. III, 865*)

Benares, Aug. 27 (now Varanasi) **VII**

An earthquake happened here on the 26th instant, the particulars of which are as under. About 23:00 26 Aug. a shock was felt which continued for 5–6 s. This was followed by another, though not very violent, of 7 or 8 s duration. On going out to the verandah I found several of my friends and the domestics in a state of alarm. After a lapse of 5 or 6 min I returned to my chamber, but no sooner had I entered it, than a more severe violent shock was felt. This lasted 2.5 or 3 min and the earth ceased to shake gradually and imperceptibly.

(*The India Gazette, Calcutta, Wednesday, 4 Sept. 1833, III, 865, reprinted in Bengal Hurkaru and Chronicle XII(58), New series, Thu., 5 Sept. 1833*)

Benares **VII**

A native letter from Benares mentions that three shocks of an earthquake were felt at that station. It, however, appears to have most severely been experienced at Boglepore as detailed in a letter in the John Bull. The first shock was felt at 11:30, the undulations from the NW passing to the east. An hour after, a second shock was succeeded, more severe than the first, and afterwards six partial shocks ending at 2:30. The last one was accompanied by a sound resembling the rushing of a stream of water. A constant tremor of the ground was felt from 12 O'clock until daybreak.

(*The India Gazette, Calcutta, 26 Sept. 1833, III, 884*)

Berhampore, 4 Oct. 1833 aftershock

This morning, at a well known village called the Cuttereah, about 10 miles below Berhampore, while we were out for a walk we distinctly felt the shock of an earthquake. It came about 07:45 and lasted full 5 min; its direction was from north to south: the shock was slight but distinctly felt so that you see earthquakes are becoming more common now.

(*In Bengal Harkuru, Wed, 9 Oct., printed from Indian Register*)

Bhaugulpore, 27 Aug. 1833 **VII**

About a quarter past 11 last night, I was roused from my bed by a violent shaking of it, so much so that at first I thought someone was actually moving the bed about. I jumped out and saw the oil in the lamp moving about in tremulous motion which lasted about 3 or 4 min. I waited a little and got my wife and family on the verandah, when we plainly felt the chairs move and the doors rocking about. This lasted nearly 8 min, and between this and daylight three other shocks were felt, but none equal to the first one; a rumbling kind of noise was heard, but not very loud. In the year 1806 I felt an earthquake at Bencoolen as severe as this, but never before in this country.

(*The India Gazette, 1833, Tues., Sept. 3, III, 864*)

Subsequent commentary on above

Sir, The account contained in your paper relative to the late earthquake at Bhaugulpore is indeed terrific. The shock of 8 min duration must have been confined alone to your correspondents premises. His house must be evidently built of something more substantial than pucka materials, else how is it possible that it stood. Your correspondent adds that he felt the shock at Bencoolen in the year 1806 as severe. Judging from this, he must be an old stager; consequently, when age and fear act as existing agents, some allowances are admissible. I remain your obedient servant. – Anti Veteran.

(*The India Gazette, Calcutta, Wed., 18 Sept. 1833, III, 877*)

Bhaugulpore, 4 Oct., 08:10 aftershock

10 s of shaking described at length from the shock of the 4 Sept. 08:10. Birds disturbed. A mud hut about the neighbourhood had fallen, and an aged man, who could not get out in time, was buried under the ruins, and was taken out dead afterwards.

(*The India Gazette, Calcutta, Fri., 11 Oct. 1833, III, 897*)

Buldeeah – Inundation and Earthquake, 31 Aug. 1833 (possibly Batia, viz. Plate 12, ref. 47).

A letter from a correspondent at Buldeah dated the 31st ultimo states that the place had been visited by the most destructive inundation of the Kumlah river ever known. The river was increased by torrents from the Nipaul mountains so suddenly that hapless beings were drowned asleep in their houses at night. Cattle were swept off, and immense trees of the forest came rolling down the rapid stream that had undermined or cut away the banks on which they had flourished for many years in perfect safety. The earthquake of the 26th and 27th was also felt very powerfully at the same place.

(*The Bengal Harkuru and Chronicle XII(67), New Series, LXIV, 5768, Old Series, 16 September 1833*)

Buxar, 27 Aug. **VII**

Last night several very smart shocks of an earthquake were felt. The first decided one took place at 23:20, but the smartest of all at 23:55, each continuing for several minutes in a direction from north to south. The latter was so severe as to cause the members of all families resident here to run into the open air to secure themselves against the danger which seemed about to overwhelm them. The motion was so strong as to upset several wall shades and glasses in one of the bungalows here, and the general rattling of doors and windows, and the creaking of rafters were really most alarming.

The undulating motion of the ground was so great as to cause horses that were asleep to start up suddenly with the mark of alarm, and the birds roosted upon the trees to rise into the air, uttering wild cries.

Strange to say, the violence of the shock seemed a good deal bounded by the river, as at Koruntadhee, immediately opposite to Buxar, the motion of the earth was so little felt as hardly to awaken a single individual, whereas on the right bank of the river, all were obliged on the instant to evacuate their habitation and run into the open air.

The effect on the water about 56 miles above Buxar, as described by me, an officer luogoo'ed in his boat, was that of a rocking motion, as if occasioned by a heavy swell consequent on an east wind.

I am glad to be enabled to add that no serious injury occurred, unless one or two cracks in an Indigo factory bungalow, a few koss off, may be accounted as such. I have since heard that the shocks were most severely felt at Bhojepore, 15 miles distant in the east direction.

(*The India Gazette, Calcutta, Monday, Sept. 9 1833, III, 869*)

[Note: A kos is 3000 paces or approximately 2.3 km (p. 198 in Phillimore, R. H., Historical records of the Survey of India, Vol. II, 1800–1815, 477 pp.). The above account, though taken at face value by the commentator in *Mofussul Ukbar of 21 Sept. 1833*, and repeated by Mallet, 1852, is ridiculed in a reply in the Bengal Harkuru, Fri., 20 Sept. 1833, an extract of which is reproduced below.]

Dear Sir, We have all been a good deal amused here, with the account given in your paper of the 6th instant, of the earthquake at Buxar. 'It would certainly be important if true' – but there's the rub! Why! Mr Editor, your ordinary discrimination must surely have been in abeyance when you admitted so bright an idea as that 'the effects of the shocks could even appear to be bounded by the river'.

The fact is that the motion of the earth was felt just as severely at Koruntdhee as at Buxar on the right bank of the river, with this small difference, that your gallant and philosophical correspondent having previously confined a considerable time in his boat going down the river, sacrificed a little too freely to the Jolly God, at a certain glorification party on the night in question ... Servants, syces, etc., rushed out with one accord in the greatest state of alarm. In short, it was impossible not to have felt it. Yours truly, VINDEX. Buxar, 13 Sept. 1833.

(*Bengal Harkuru, Fri., 20 Sept. 1833*)

Calcutta

Sir, The correspondents who have favoured you with accounts of the earthquake that was lately felt in various parts of Hindoostan state in vague terms the direction which the earthquake took in its progress. Do they mean to say that the direction was north and south or from north to south (for these expressions convey different ideas to the mind). If the latter, what means have they of ascertaining the fact? Your obedient servant, B.

(*The India Gazette, Calcutta, Tue., 12 Sept. 1833, III, 872*)

Calcutta **V**

Two shocks of an earthquake were felt on Monday night: the first one occurred as is shown by its stopping an astronomical clock precisely at that time 23:35, and was comparatively slight and of short duration: the second one was far the most violent we ever remember to have experienced – the doors rattled, as if shaken by a storm of wind, the bed in which we slept was heavily for more than a minute, from north to south, with a gentler undulation afterwards in an easterly and westerly direction: the punkahs, the wall shades, the tables and chairs were violently agitated; and this second shock endured we should think altogether, for upwards of 2 min. It came on at about 23:56. The natives felt it in their huts and blew their horns as usual, and the crows and adjutants were disturbed by it and made a great noise, flying off from the houses and trees where they had taken roost.

(*Bengal Harkuru and Chronicle, XII(51), New Series, Wed., 28 Aug. 1833*)

Calcutta, 27 Aug. 1833 **V**

A slight shock of an earthquake was felt in Calcutta last night (Monday) at about 11:35. A stronger shock was felt at exactly 23:57.

Subsequent commentary on the above

Sir, you must be a sound sleeper from your editorial of this morning about a slight shock of an earthquake felt last night. I have witnessed many; and, except for that experienced in Lord Lake's camp on the night preceding the storm of Ally Ghur (note: 1 Sept. 1803, Kumaon earthquake). I remember nothing of the kind comparable to the lengthened rocking of my couch last night in the upper rooms of the Bengal club. I really expected to see the pillars of the verandah, if not the building itself, come down, and waited with tranquillity for my 'disruption of the visible world', not the oft recurring wish to behold what is beyond.

(*Cleophas. India Gazette, Calcutta, Wed., 28 Aug. 1833, III, 859*)

Chandernagore **VI**

The following is an extract of a letter from Chandernagore: 'In the night of the 26th–27th two shocks of an earthquake were felt in Chandernagore; the second one, much more violent, continued for about 3 min and commenced about half past 12 (*vers minuit et demi*). The attic of a house was thrown down, but no other accident occurred.' We suspect that our correspondent has mistaken the time of the last shock.

(*Bengal Harkaru, Thu., 29 Aug. 1833, XII, No. 52, New Series*)

Chuprah, 29 Aug. **VIII–IX**

In the night of the 26th instant we were thrown into great consternation by the most awful shocks of an earthquake, perhaps ever felt in this part of the world. The first was at about 5 O'clock in the evening and slightly felt. At midnight I was suddenly aroused from sleep by a tremendous rumbling noise and violent rocking of the house, which continued for some time. I remained in the verandah thinking that all was over, when in about a quarter of an hour another shock began, first with a tremulous motion and a noise like thunder, followed immediately by a very alarming and awful agitation of the earth, which vibrated dreadfully in a horizontal direction north and south.

We all rushed downstairs away from the house and continued in an open building until daylight, and as I did not sleep again, I observed that slight shocks continued till 8 in the morning; indeed, a slight agitation was felt at intervals throughout the day. Large pieces of the cornice and much of the plaster fell from the walls of the rooms.

All the people ran from their houses and several native dwellings were destroyed. I have since heard that three people were killed by the falling of their houses. As such severe shocks are unusual in this part, you will probably learn further accounts by the papers. Two miles from this the earth opened and a chasm of considerable depth formed. In consequence of the injury the Court house has sustained, Cutchery is now held in a tent.

(*The India Gazette, Calcutta, Fri., 6 Sept. 1833, III, 867, reprinted in Bengal Hurkaru, XII(60), Sat., 7 Sept. 1833*)

Comercolly (Khumarkhali, 23°52′, 89°11′) 27 Aug. 1833 **VI**

Two very smart shocks of an earthquake were felt here last night

(*The India Gazette, Tues., 3 Sept. 1833, III, 864*)

Delhi, 27 Aug. **V–VI**

An earthquake was experienced at Delhi at a quarter before 1 AM on the 27th instant which lasted 4.5–5 min; the shocks (by undulations) were rapid and distinct and their direction was from SE to NW.

(*Delhi Gazette, 31 Aug., first reprinted in the 11 Sept. issue of India Gazette 871 and again Wed., 18 Sept. 1833, III, 877*)

Futtigurh, Sept. 28 1833 **V**

Your 'excellent earthquake' was felt here very sensibly, I was sound asleep when it occurred, but the clattering of the doors and windows soon awoke me and made me leap out of bed. From the harmonious cawing of the crows and the sweet carolling of the minahs (?), I soon returned to my nest. A neighbouring house was much shaken.

We were subsequently visited by a meteor so brilliant that I was again woken from a sound sleep.

(*India Gazette, Calcutta, 10 Oct. 1833*)

Goruckpore, 27 Aug. 1833 **VII**

The first shock was felt at about 7 pm yesterday and was slight. At about 23:20 there was a severe shock, which lasted nearly 2 min; this shock awoke without exception, that I have heard, every European resident at the station, many of whom ran out of their houses, fearing that they would tumble about their ears, and really the shaking was so violent that it felt exactly as if the walls must be thrown down. After this shock passed away there was a continued vibration of the ground for about 20 min, at a few minutes before 12 another shock, much more violent and of longer continuance than the last mentioned, was felt, when open doors were slammed to, the jilmills rattled and the violence of the shaking was such that it seemed impossible for any fabric to stand it. The city was in an uproar, all the natives ran out into the street and continued out the remainder of the night. There was an almost constant vibration felt until nearly 6 O'clock this morning and three shocks, much less severe than the two between 11 and 12, at intervals between 12 and 6; since then all has been quiet. One or two houses in the town were shaken down but the only damage I have heard in the cantonments is the cracking in two places from top to bottom of the outer verandah wall of the civilians bungalows. It was a beautiful moonlit night without cloud or a breath of wind.

The shocks appeared to run from east to west – it was not an undulating motion such as I have at former occasions felt, but the sensation to a person in bed was more like what be felt if the bed were seized by several hands and shaken violently backwards and forwards. The natives here have scarcely recovered from the fears yet; they say they never experienced anything of the sort before, and it seems surprising that no houses were injured more materially, considering the violence with which they seemed to be shaken.

The river was exceedingly high and all the jeels and low ground with which it can communicate more extensively filled and flooded than they have been since the great flood in 1823, though the local rains have not been excessive; indeed, there has scarcely been an average fall and the weather has been delightful.

(*Bengal Hurkaru and Chronicle, XII(59), New Series, Mon., 6 Sept. 1833*)

Goruckpore (18 October aftershock) **VIII**

A correspondent writing from Goruckpore says that the station was visited on 18 Oct. 04:40 by 'a most tremendous earthquake' which came from east to west and lasted nearly 1 min. He adds that it was much severer than that which was experienced on the 26th last. It is singular that we felt nothing of the kind here, nor have heard anything of the earthquake from other quarters.

(*The India Gazette, Calcutta, Sat., 26 Oct. 1833, III, 910, reprinted from the Englishman*)

Hatrass (now Hathras) **V**

A considerable shock of an earthquake was felt at Hatrass. It appeared to last a minute and it seemed to be from NW to SE.

(*The India Gazette, Calcutta, Tue., 12 Sept. 1833, III, 872, reprinted from the Mofussal Ukhbar, 31 Aug. 1833*)

Jionpor 10 Oct. 1833 aftershock (Jaunpur) **V**

I lose no time in telling you of another earthquake, which was felt here yesterday at or about 8 in the morning; the shock was sudden and smart and lasted a few seconds. It shook all the jillmills and doors. Jionpor, 11 Oct. 1833.

(*The India Gazette, Calcutta, Sat., 26 Oct. 1833, III, 910*)

Kathmandu, Nepal, 27 Aug. **IX–X**

Have you ever experienced the sensation from the shock of an earthquake? If not, pray earnestly you may be spared a practical knowledge of this huge and incomprehensible phenomenon. Last evening about 18:00 we had in the residency grounds, and throughout the valley, a severe shock of an earthquake. I was in the house at the time. Doors, jillmills, chairs and tables commenced shaking most violently, then the house began to move, and when I got outside, the rumbling noise of cannon driving over a drawbridge was most distinctly audible, and gave the idea of vast bodies moving at full speed close to underground. The trees and shrubs moved from their roots with each undulation of the ground, and I could distinctly see my house reeling in unison with the trees around it. I pitched a tent to have a place of refuge against a repetition of the business. However, bed time came and we all save H. went to bed; but at 10:45 another and more violent shock than the first one obliged us to turn out, and as I rushed out of the house instantly, I had a noble view of the effects of this one. The ground moved in a long and even swell for about 30 or 40 s, and the doors and windows of our houses shook most vehemently. The whole people of the place set up a shout – the Hindoos called on Ranjee, and the faithful uttered long and deep prayers to God and his prophet – all were in confusion and amazement – when at 11:02 we heard the rumbling of a distant thunder and then a shaking and reeling of the ground that was quite fearful, accompanied by the noise of falling tiles, accompanied by a loud shout from the city which was quite audible here at a distance of a mile. I cannot tell you what were my own sensations on feeling the solid earth reel as a drunken man.

From this time till half past twelve we had six severe shocks and they continued until six in the morning. The residency is sadly injured – the whole of the upper story is nearly gone. R's house is rendered uninhabitable, and even mine, which is a pucka one of only one story, has got some awkward cracks in many parts of it. An elephant house, cook rooms and other small houses are much injured, and we want but a few more shocks to bring our houses down. No lives have been lost on the residency grounds, but it is otherwise at Catmandoo, Patna [*Patan/typo*] and other towns in the valley. The loss of life is not yet ascertained, but they say 25 in the former and 30 in the latter town. In both towns several houses have been levelled with the ground, and a large temple of Juggernauth, which had been formerly injured by lightning but was near an immense mass of pucka buildings, came down this morning by the run, and presents nought now by a mountain of bricks. The houses of many chiefs in the city have suffered much, and two pillars, each about 100 ft high, built by the minister and his nephew some years since, are now no more. The city is at present quite deserted, the Rajah having ordered every man to vacate his house and take to the fields for two days. Fair ones, who till last night had never seen the light of heaven, save through a purdah, were bundled *sans ceremonie* into the open air, and all in all it is a fearful state of things. I trust it is over for the present, but we are in doubts, for the last great earthquake here in 1829 continued for many days, and even at this sitting a slight shaking of the ground is perceptible.

Letters continue to arrive from all parts of the country especially from the north and west, giving accounts of the effects of the earthquake, some of which seem to have been of most serious character, involving destruction of life and property. It does not appear, however, as far as our information at present extends, that the calamity has been so heavy as was a similar visitation in 1829, while in comparison with the earthquake in 1762 it has been extremely mild and innocuous. On the last occasion the rivers and tanks throughout the country were strongly affected – the waters rushing with violence and rising in some places ten feet perpendicularly, and continuing in agitation for a quarter of an hour at a time. At Dacca 500 people were killed, at Chittagong the earth opened in many places, quantities of water gushed out, great chasms remained unclosed, reports underground were heard, and many buildings were thrown down. Near Luckipore, a circuit of land, about 15 miles in circumference, was swallowed up, and all the inhabitants and cattle perished. Nothing equal to this has occurred till now.

(*The India Gazette, Calcutta, Monday, 9 Sept. 1833, III, 869, reprinted from John Bull*)

Katmandoo, 30 Aug. **X**

On the 26th at 11 O'clock PM we had a serious shock of an earthquake, which is calculated to have destroyed 8000–10,000 houses and 600–800 souls in the several towns of the valley, and to have done extreme mischief beyond its limits eastward. The direction of the motion was NE and SW and its character, undulatory, with a very long swell and roll. The great shock was preceded, 5 h earlier, by a lesser one, and has been followed by innumerable trivial vibrations which have not yet ceased.

(*The India Gazette, Calcutta, Tue., 12 Sept. 1833, III, 872, reprinted from the Bengal Hurkaru Wed., 11 Sept. XII(63), New Series*)

Katmandoo, 30 Aug.

Further accounts have been received from Nepal describing the terrible effects of the Earthquake of August at Lassa, where the destruction of lives and buildings has been the greatest. The particulars, we understand, will be given in the *Journal of the Asiatic Society*. A late number of the *Mofussal Ukhbar* has a long article on the last earthquake in which reports are given for a great many places, showing pretty accurately the extent of its influences. The writer concurs with the author of the article in the *Journal of Science* in tracing the seat of the disturbing cause to some unknown spot among the Himalayan mountains. A fact is mentioned in the Chupra Report which merits the attention of our geologists if true, but which we should not hesitate to believe without further details, namely that the earth opened, and a chasm of some depth was formed about two miles from that place.

(*The India Gazette, Calcutta, Sat., 6 Oct. 1833, III, 892, reprinted from the Calcutta Courier*)

[Note: The J. Asiatic account is Prinsep's p. 438 in the August edition, and the Mofussal Ukhbar reference may refer to the 14 Sept. edition above that may contain more details than printed in the India Gazette. The comment about Chupra above refers to the entry in the 6 Sept. India gazette, III, 867, last paragraph on Chupra. Campbell is adamant that no damage occurred at Lhasa, and may have been prompted by the above news report to be emphatic about this in the November edition of J. Asiatic Soc.]

Lucknow, 27 Aug. **VII**

It may be interesting to your readers to learn that last night (the 26th instant) the city of Lucknow was visited by four shocks of an earthquake, the first about sunset, the others between 11 and 12 PM. The tremulous motion of two of the shocks resembled the motion felt in a steam vessel; the rocking caused the beams of the houses to creak, the lamps to swing to and fro, and in one house fragments of cornice to fall. In the stiliness of midnight, superadded to the knowledge that

the scourge of Cholera Morbus was now fearfully active throughout the city, there was something very awful in the sensation, and as the earth rocked and the walls moved, I almost expected to hear the crash of some falling building.

(The India Gazette, Calcutta, Tue., 12 Sept. 1833, III, 872, reprinted from the Mofussal Ukhbar, 31 Aug. 1833)

Lucknow **VII**

On the night of 26 Aug. an earthquake was felt here, some people felt several shocks, one about sunset. I felt only the last one, which must have been the most severe. It awakened all out of our sleep at half past eleven at night and lasted about 2 or 3 min. A servant of mine who thought the house was coming down ran across the compound to his own house, and fell down three times on the way, the ground shook so much.

(The India Gazette, Calcutta, Wed., 18 Sept. 1833, III, 877 and in Bengal Hurkaru, Fri., 20 Sept.)

Meerutt **V**

Earthquake. Meerutt was visited on the night of Friday last by the smart shock of an earthquake. It continued for about 15 s, but was not severe enough to do any mischief. We have received further accounts of damage done by the late earthquake in the country immediately above Benares. The tract lying between the Ganges and the hills seems to have been most severely visited; details have, however, been published in the Calcutta newspapers.

(The India Gazette, Calcutta, 10 Oct. 1833, III, 896 and in the Beng. Hurkaru and Chronicle Mon., 7 Oct. 5786, Old Series, reprinted from Meerutt Observer, 26 Sept.)

Malda, 4 Oct. 1833 aftershock **V**

We had a severe shock of earthquake yesterday at 08:10, which set the punkah swinging and venetians rattling in a fearful style. Malda, 5 Oct. 1833.

(The India Gazette, Calcutta, 10 Oct. 1833, III, 896)

Monghyr, 27 Aug. 1833 **VIII–IX**

Sir, Five shocks of an earthquake were experienced last night and this morning at this station. The first one occurred 5 min after sunset, the second one at 11:25, the third one at midnight, the fourth and the fifth ones at 03:45.

The second and the third ones were most severe and caused considerable damage to the houses in the town, of which about 34 were more or less injured by the falling of balustrade and balconies or cracking and separation of the sidewalls and beams.

Mr. Hamilton's house suffered severely, and the beams of the verandah all loosened from their sockets in the walls, and many feet of plaster cracked and torn off, the walls having been also much shattered.

At the third shock the inmates of my house all alit and passed the rest of the night in a temporary covering made with bamboos and a suttringee. The lamps were all extinguished by the trembling of the walls, the clocks stopped, and the glasses of the sideboard knocked together and made, with the agitation of the jhilmills and rumbling underground which accompanied the earthquake, a frightening noise.

The direction of the shocks appeared to take was from the NE to the SW. This circumstance was manifest from the noise made by the natives as the shock passed along. The people continued shouting and drumming the whole night, I suppose to propitiate the causes of this awful visitation.

I am unable to speak with any certainty of the duration of the shocks, but from what I have heard and from my own recollection, the first one appeared to last about 30 s, the second and the third ones about a minute each, after which there appeared to be a trembling motion in the earth for about half an hour, and the glasses in the sideboard continued during that time to shake.

Colonel Bird's house sustained considerable damage and his family was obliged to leave it altogether. They took refuge during the night in some tent pitched during the day to be aired.

Considerable portions of the ramparts of the fort fell down; I observed two places in my mornings ride to the extent of about 20 yards in breadth and 30 in height. These walls are made of brick and mortar cased with stone. I heard that there were two other parts of the wall that had been thrown down.

Mr. Page's house and Captain Watson's were also, I heard, much damaged. In my own the outer walls on the East and West sides cracked in several places, the fissures going quite through the walls.

There was no particular change in the river that I am aware of, but the shocks were distinctly felt by people in boats. One gentleman I met this morning said he thought at the time that a violent storm was coming on from the agitation of the water and the motion of the boat, but finding the true cause from the people on the shore he deemed it prudent to leave his boat.

So general was the consternation caused by this unusual event that nearly all the Europeans and other inhabitants of pucka houses resident in and about the fort left their domiciles and sought refuge in the maidan.

Yesterday was a particularly hot day, and the evening sultry and unpleasant. Today we have much the same kind of weather, but the horizon is surrounded by heavy lead-coloured clouds, and there is lightning in all directions.

Since the last shock above alluded to, partial and slight tremblings of the earth have been experienced the whole day.

The commissioner's house, situated on the highest spot at the station, did not suffer in the least; the earthquake, however, was severely felt there. Since writing the above another slight shock occurred at half past none P.M. Yours faithfully, A Resident.

(The India Gazette, Calcutta, Thur., 5 Sept. 1833, III, 866, reprinted in Bengal Hurkaru and Chronicle XII (59, New series), Mon., 9 Sept. 1833)

Monghyr, 27 Aug. **VII**

This station was visited by an awful calamity: a constant succession (about 30) of awful shocks from 17:00 yesterday until 08:30 today, of which some were so severe as to destroy several fine houses and materially injure others. The whole population of Monghyr spent the night in the open air, and we much feared that this place is the focus of volcanic matter, and that the present is but the forerunner of a still more severe visitation.

(Bengal Hurkaru and Chronicle XII(58) New series), Thu., 5 Sept. 1833, reprinted in The India Gazette, Calcutta, Fri., 6 Sept., III, 867)

Monghyr, 27 Aug. 1833 **VIII–IX**

We were last night visited by one of the most awful calamities that has ever been known in this place. A succession of seven shocks of earthquake from 17:00 till 08:30 this morning, in which many fine houses were injured and some totally destroyed; amongst the latter is Mr Graham's fine house, which was rented on a lease to Lt. Col. Bird. It was truly awful to hear the cries of the people and crash of houses falling, the ringing of bells, the beating of tom-toms, the Musselmans at prayers, and all the populations of Monghyr, of every description, out in the open air; the ground every 5 min shaking and trembling in a frightful manner. Mr B. and family were obliged to fly

for their lives, and come over to mine; several others did the same and amidst the horror and dismay of all around we spent the night. There were upwards of 25 shocks during the night, and 5 after day break.

(The India Gazette, Calcutta, Mon., 9 Sept. 1833, III, 869, reprinted from Bengal Hurkaru, XII(60), Sat., 7 Sept. 1833)

Monghyr

A letter from Monghyr mentions that shocks of earthquake were felt there up to the 29th ultimo and a letter from Mullye which we publish shows that they were felt there up to the 31st [August] inclusive. We are much obliged to correspondents who have favoured us with such accounts. They will doubtless be of use, as has been suggested, in preparing a connected statement of all the phenomena.

(The India Gazette, Calcutta, Wed., 11 Sept. 1833, III, 871)

Monghyr, 2 Sept.–Mon. 26 Aug. Earthquake **VIII**

Several pucka houses in the town sustained injury, and all are more or less cracked. I observe that the houses most damaged are those that have new buildings attached to the old – the same is the case with houses in the fort. I hear that all the clocks at the station stopped at the second, or severe shock.

(Calcutta Courier, 11 Sept., reprinted in Bombay Courier, XLI, 2277)

Monghyr, 12 Sept. 1833

You would no doubt be surprised at my last, how much more you will be astonished to learn that the earthquake mentioned in that has not entirely left us, we have had numerous shocks since the first one, but of no great magnitude, and perhaps they have added more fear than that to real injury, but certain it is that the earth has never been for even 5 min perfectly steady since the 26th of last month, and from the continued tremulous motion, many are fearful that it will end in some dreadful catastrophe.

(Bengal Hurkaru and Chronicle, Sat., 21 Sept.)

Morshedabad zilla, 30 Aug. 1833 **V**

We were visited on the 26th with several smart shocks of earthquake; the first one about 19:00 was but slight, but two others about midnight, within about 10 min of each other, shook the house, awakened the birds and caused the hanging lamps to vibrate for 20 min after – so severe was the last shock that I feared the house would be injured. The earth trembled so violently that it was difficult to keep one's leg. The natives in their frail tenements felt it severely for the neighbouring villages resound with their 'Hurribole, Hurribole'.

(Reprinted in Bengal Hurkaru and Chronicle, XII, 59 New Series, Mon., 6 Sept. 1833, from John Bull)

Mullye, 27 Aug. **V**

You will doubtless have numerous accounts of the earthquake which occurred last night at this station. The first shock, which was very slight, took place at 18:30. At 23:15 we had a very severe shock, and about 23:45 another one which continued whilst I deliberately counted 350, for I had no watch by me to time the shocks; some seconds must have also elapsed previous to my commencing counting. From this period until 03:15 we had a succession of slight shocks, in all 16 in number, and at 06:00 a seventeenth. They ran in a direction from north to south and the vibrations occasioned by the second and third were particularly violent, very nearly, if not quite as violent as the great earthquake in 1803. Meteorite at 3 am. Your obedient servant.

(The India Gazette, Calcutta, Fri., 6 Sept. 1833, III, 867, reprinted in Bengal Hurkaru, XII(60), Sat. 7 Sept. 1833)

Mullye, 30 Aug. 1833

Since the first night of repeated shocks at Mullye, we have experienced several slighter successions, on each succeeding day and night, up to this date.

(The India Gazette, Calcutta, 1833, Mon., 9 Sept. 1833, III, 869, reprinted from the Bengal Hurkaru. Tue. 10 Sept. 1833, XII(62), New Series)

Mullye, 1 Sept. 1833 **V**

I mentioned that we had no less than 17 shocks of earthquake at this station during the night of 26th up to 6 A.M. ultimo. In the course of that day two other slight ones occurred and two more from 9 to 10 P.M. On the 28th at 16:15 we had a very smart shock, the vibrations of which lasted for about a minute. On the 30th at 11:30 a slight one, at 01:00 this morning one shock, at 06:00 another shock. Numerous other tremors of the earth have been experienced, but we have become so accustomed to the sensation, that it is difficult to divest ourselves from the idea of the shaking, and it is only when assisted by the other senses, and that when we perceive the flapping of doors and moving of punkahs, etc., that we are satisfied we are not under a delusion.

(The India Gazette, Calcutta, Tue., 10 Sept. 1833, III, 870)

North Tirhoot (North of Muzaffarpur), 27 Aug. 1833 **VIII**

This part of Tirhoot was visited at 18:00 yesterday by a severe shock of an earthquake, which began with a motion from east to west and which was repeated at 11 and 12 with double violence continuing throughout the night until 6 this morning. The two shocks in the middle of the night were most awful, lasting I should think, each time, more than a minute. The bungalow in which I reside was violently shaken, rocking to and fro, the doors and wall shades were equally agitated, and the walls were rent in many places. Had the successions towards morning been equally violent, I fear my bungalow would have been in ruins.

The walls of my press house and drying house were opened in several parts and the water in the reservoir 4 ft deep and 3 ft from the surface was so much agitated as to fly over each side. At one time I thought all was going to rack and ruin.

I have been in India for many years, and I have never experienced anything half so awful as the shocks during the night. In 1818 [probably Cutch earthquake 1919] if I remember right, a severe shock was felt in Calcutta, but nothing equal to the one of last night.

(The India Gazette, Calcutta, Mon., 9 Sept. 1833, III, 869)

Nusseerabad (Nasirabad, 24°26′, 90°) **I**

'No mention is made of the earthquake.'

(The India Gazette, Calcutta, 26 Sept. 1833, III, 884)

Patna, 28 Aug. **VIII**

But now we have recovered from the horror of the event I have to communicate of an earthquake which has alarmed us exceedingly. Its first access was 18:30 on the evening of Monday, and I distinctly felt the shock, but as I was not supported by any observation of similar tendency, I began to think it must have been fancy. About 23:30 I

was awakened by a horrid rattling and shocks apparently from east to west with a rumbling noise, the servants running into the room in great consternation, when the whole truth flashed upon my mind. This was the first of two very severe shocks, the last of which occurred at midnight, and the oldest inhabitants here say that a severer one they never remembered. It is quite impossible to describe the horrors of these two shocks. I have faced a good deal of nature's wildest wrath, the wrath of the ocean and storm and tempest, and in both I have seen the hand of God in death – and other visitations, but these were indeed 'the terrors of the Lord'. Several houses at Bankipore were cracked by the earthquake, the Magistrate's Kutchery being one amongst them, and in the city, the opium godown shared the same fate as well as the great gateway built by Maharajah Metre Jeyt Sing, while many of the pucka native dwellings have bodily fallen. At Dinapore the former Brigade Major's house was split and also some houses between Patna and Diggah.

Eighteen shocks were counted by some, but they appeared continuous. The last one was at 08:00. Almost all the inhabitants left their houses during the continuance of the shocks and remained out the whole night.

(The India Gazette, Calcutta, Monday, 9 Sept. 1833, III, 869, reprinted from Bengal Hurkaru, XII(60), Sat., 7 Sept. 1833)

Patna, 27 Aug. 1833 **VIII**

Last night at about 23:00 we had a most severe shock of an earthquake the like of which I have not witnessed or heard of; it came on at four different times, and the shock was so very great that the whole of Patna appeared as if floating and a great number of houses fell, and various sorts of injury sustained. Rajah Khan Behadoor's Stable fell by which seven of his horses died.

The outer wall of Captain Elliot's house was also levelled with the ground. Mr E. D'Costa's house also cracked at several places and a few of his wall shades fell to the ground. Many of the gentries suffered considerable injury to their property. The alarm was so great that the inhabitants left their house and stood out in the compound with their family the best part of the night. The shock was most dreadful and alarming, we expected every moment to be our last.

(The Bengal Hurkaru and Chronicle, New Series XII, 57, New Series, 4 Sept. 1833, Calcutta)

Purneah, 27 Aug. (now Purnea) **IX**

This place I am sorry to say has been visited by several of the severest shocks of earthquake within living memory of the oldest inhabitant. The following is as correct account of it as I was able to preserve:

1st	shock at 17:30, 26 Aug. 1833, rather slight	
2nd	23:00	severe
3rd	23:18	most violent
4th	23:28	slight
5th	23:32	slight
6th	23:46	slight
7th	01:00	27 Aug. 1833 slight
8th	03:25	slight
9th	05:52	slight
10th	08:00	rather severe

So great was the undulation that birds of every description were driven from their nests, men were scarce able to keep their feet, and the cattle running about wild with fright. During this severe shock, several old buildings in the town tottered to pieces, and part of one of the stations actually partially sank. A heavy cornice of a newly erected wall in a gentleman's bungalow came bodily down and literally smashed a table and some chairs to pieces: every building suffered more or less. A recently vacated upper-roomed house had not an arch left; large masses of bricks had fallen out; the roof in many places slipped and the beams and burgahs separated by an inch or more; the outhouses in this compound became a heap of ruin.

I am led to infer from the stoppage of a clock and other testimony that the undulation came from southward and traversed east. During the 3rd and severe shock a heavy rain fell to the south accompanied by much wind. I have no time to particularize further. You will be glad to hear that no lives were lost, at least, I have heard of none, and I trust I shall not.

(The India Gazette, Calcutta, Fri., 6 Sept. 1833, III, 867, reprinted from the Bengal Hurkaru and Chronicle XII(58), New Series, Thu., 5 Sept. 1833)

Purneah, 27 Aug. 1833 **VII**

The earthquake seems to have been felt there as elsewhere on the night of the 26th: no less than 14 or 15 shocks were experienced – two of them which took place between 23:00 and 24:00 were very severe, and appear to have done a good deal of injury to several houses of the station.

(Reprinted in Bengal Hurkaru and Chronicle, XII, 59 New Series, Mon., 6 Sept. 1833, from John Bull)

Purneah, 11 Oct. 1833 (aftershock)

The 4th Oct. shock at this place though not severe was sufficiently so to be duly appreciated. Upon the testimony of a gentleman whose premises lay not 5 miles of the Nepal boundary, and who was at the time shooting on his grounds, I learn that immediately anterior to the shock, three distinct reports similar to the discharge of heavy artillery were distinctly heard followed by a slight and jerking vibration, which instantly assumed an extraordinary undulation, corresponding with the rise and fall of a heavy swell at sea... in the villages of Furkeah and Shanore... Yours obediently, A reader.

(Bengal Hurkaru, and Chronicle, Sat., 16 Oct. XII(102).)

Rotas hills, 28 Aug. 1833 **IV**

At 23:10 on 26th instant, just about falling asleep, felt my bed shaking and the roof the bungalow creaking; an earthquake says I, and got out of bed immediately, and entirely out of the house? Not a soul near but a chokedar who seemed just to have awoke. I asked him if he felt the earthquake shake, he said he thought the house shook, but it was owing to wind, of which there was only a light air from the NW – a pretty bright moon; this shock lasted not more than a minute I believe; went to bed again, and lay watching, when a second rocking came on, and at 23:28 it came on again, sure enough much more severe and of much longer duration – I really think of not less than 2 or 3 min. I lay quiet some seconds to be certain it was not the effect of fancy, but it became too palpable; the roof of the house was creaking, and the lamp swinging 2.5–3 in. each way.

(Bengal Hurkaru and Chronicle, XII(59), New Series, Mon., 6 Sept. 1833)

Rungpore, 27 Aug. 1833 **V**

We had last night three earthquakes. The first a little after sunset, very short. The second at half past eleven, very strong and it lasted about thirty seconds or more. The third was still more severe and of forty seconds duration. Unlike those in July they began gently and went on fearfully increasing to near the end.

(The India Gazette, Tue., 3 Sept. III, 864)

Tirhoot (Muzaffarpur region), 5 Sept. 1833 IV (?X)

Then comes the 26th – the awful earthquake of which you no doubt ere this had better accounts than I can pretend to give; however, with your permission I will state how I felt myself. At 6:30 there was a slight shock; again at a quarter past eleven I was aroused from my slumbers by a violent shaking of the bed under me, and immediately conjectured what was going on, and made a precipitate retreat to the *maidan* fully expecting the house would be down before I was out of danger. This shock passed over in a short time, I should say it lasted a minute, and I again turned in, and was just on the eve of falling asleep, when came another edition far more severe than the preceding, which again sent me flying or rather staggering, for I certainly found it difficult to retain my gravity. To give you some idea of its tremendous force, I will merely mention the circumstances of my reservoir which was full, having had the water thrown out to the depth of two feet and a half North and South, by the vibration. The damage done to the Pucka buildings and bungalows in the district is considerable, but no doubt still more so at Patna and such like large cities, where the houses are built of frail materials from economy, and shew being the chief object without regard to stability. We have repeated shocks (slight ones) every day since; a very perceptible one yesterday at 3 P.M. and what the end of it all may be a subject of speculation. A gentleman on the border of the Nepaul Terrai writes, that he expects on getting up some of these mornings to see a splendid volcano burning on the top of the mountains. What an addition this would be to their already majestic appearance!

(The India Gazette, Calcutta, 1833, Wed., 18 Sept.1833, III, 877 reprinted from the Bengal Hurkaru)

Tirhoot (Muzaffarpur region) ?IX

The earthquake was very grand here, and equally so all along the Terai, doing great damage of course. Accounts from Nepaul speak of it as awfully sublime at Khatmandoo, the earth rocking like a raft on a rough sea: houses, temples, walls, all tumbled about and many lives were lost. The poor people were driven out for safety into the fields, and the shocks continued at long intervals on the 27th, 28th and 29th.

(The India Gazette, Calcutta, 26 Sept. 1833, III, 884 reprinted from Mofussal Ukhbar 14 Sept.)

Prinsep, H. (ed.) *J. Asiatic Soc. of Bengal*, 2, Miscellaneous, IX, 438–439, August 1833

Note: Text assembled by the editor (H. Prinsep) from newspaper reports and A.C. (= Dr. Archibald Campbell, Assistant Surgeon to the British Legation to Kathmandu). The times are given for the second shock, not the main shock, and were used by Mallett 1852, however, Mallett was unaware of the locations of the Rothas Hills (S. of Patna), Bankura (N. of Calcutta) and Mallai. Prinsep cites Mullye as Mallai. Campbell indicates the Kathmandu time of the second shock as 22:45 which is significantly earlier than times listed from other locations. On page 564 of J. Asiatic Soc. the editor revises the Kathmandu time difference here entered as 12 min to 51 min to yield a Calcutta time of 11:36.

Earthquake of 26 Aug. IX

The daily papers have published notices of this phenomenon as observed at a great many places in the interior of India, with more or less detail, from which the following general facts may be gathered.

The direction was from NE to SW: there were three principal shocks; the first one at about 18:30, the second one at 23:30 and the third one at 23:55 (Calcutta time). In the places where it was most felt, slight and continued vibrations seem to have been experienced for the whole day following. As the time of the second vibration was accurately noted in Calcutta by the stopping of an astronomical clock, we may assume that it was the best point of comparison with the times noted at other distant points. A few of them may be classified as follows:

	Observed h:m	Diff. Long. m	Cal. time h:m	
Katmandu, Nipal, second shock	10:45	+12	10:57	Very severe, loud noise
Rungpur, second shock	11:20	–2	11:18	Many houses injured, very severe, loud noise
Monghyr, second shock	11:27	+7	11:34	Walls cracked, noise heard
Arrah, second shock	11:15	+14	11:29	Walls injured
Under Rotas Hills, second shock	11:10	+20	11:30	
Gorackpur, second shock	11:20	+19	11:39	Walls cracked, etc.
Allahabad (vague) second shock	11:00	+28	11:28?	Hollow sound from river
Bankura, second shock	11:30	+4	11:34	None such since 1814
Calcutta, second shock			11:34:48	No injury done

At Monghyr, Rungpur, Muzaffarpur, Mallai and other places within the direct line of influence, many houses were destroyed or injured, and the alarm was great. At Katmandu, the following extract of a letter from Dr A Campbell, dated the 28th instant will show that the consequences were more serious, and judging by the course of the phenomenon we may reasonably fear some dreadful catastrophe towards Lassa on the north of the great Himalayan range.

'On the evening of the 26th, at about 6 O'clock, the valley and the neighbouring hills were visited by a severe shock of an earthquake: it lasted about 40 s, and during its continuance, there was a distinct audible noise as of an ordnance passing rapidly over a drawbridge. It seemed to come from the east, and I felt that it was travelling with the speed of lightning towards the west, and just under my feet: the houses shook most violently, and trees, shrubs and the smallest plants were set in motion, not shaking but waving to and fro from their very roots. No damage was done to life or property. At 22:45 we had a similar shock in brevity and duration, and at 23:00 a most tremendous one. It commenced gradually, and increased until the houses, trees and everything on the face of the ground seemed shaken from their foundations. The earth heaved most fearfully, and when the shock was at its worst we heard the clashing of falling tiles and bricks in every direction; and to add to the impressiveness of the scene, a general shout rose from the people in all directions. The murmur of human prayers was carried audibly from the city to our grounds (a mile), and nothing could be more imposing and vast than the scene. In a dead calm the noise of a hundred cannon burst forth: full-grown trees bended in all directions, and houses reeled about

like drunken men. In our grounds no lives were lost, but in Katmandu 19 persons were buried under the ruins of their own houses, and in the towns of Bhatgaon and Patan, many more. The great shock continued for nearly a minute, and during the following hour there were six distinct and strong shocks, the ground in the intervals between scarcely, if at all, steady; and from this time till yesterday morning there were upwards of 20 distinct and sharp shocks. The loss of property has been very great, 125 houses fell in Katmandu during the night of the 26th, and nearly as many more were levelled with the ground. Up to this time, in consequence of the torrents of rain that came down, finishing the work of destruction commenced by the earthquakes, the city and town have been evacuated; men, children, and women of the *purdah*, rich and poor have been and still are on the plains about the towns. Innumerable temples were destroyed, and the very gods of them were crushed to atoms. A fine and large brick temple (100 ft high), built in imitation of the great one at Jagarnath, came down by the run early yesterday morning, and two fine pillars by Bhim Sen were demolished by the shock. All yesterday and last night we had occasional small shakes, but none equal to the great one we had on the 26th.'

A subsequenct note from the same gentleman, dated the 30th instant, gives further particulars of this disastrous event:
'We still continue to be visited by occasional shocks of earthquake, all less violent than the great one on the 26th, but sufficiently alarming. This morning when at breakfast, we had a rather sharp one: they all seem to come from the same direction; that is from the east and north-east. The places east of Katmandu have suffered most: Bhatgaon, a large town, has been almost entirely destroyed; upwards of 1000 houses have been levelled with the ground, and few have escaped serious injury. 300 souls have perished in this town (Bhatgaon) alone, and the total number of lives lost throughout the valley, as yet ascertained, is estimated at 500. The unfortunate people in many instances are in sore distress; their stores of grain being buried beneath the ruins of their late dwellings, and without money to purchase other food. The grain shops as well as all others are shut, and the people dare not return to their houses but remain without sleep or shelter in the open air, under torrents of rain. The house of Matabar Sing (a goodly modern mansion) is quite destroyed, and the large garden houses of Bhim Sen and his brother Rau Bir are rendered, for the present, untenable. Scarce a large house in Kathmandu has escaped injury. The fort at Chisopani, on the road to this from the plains, is much injured, and almost all the Government buildings have sustained great injury.' – A. C.

Campbell, A., *J. Asiatic Soc. Bengal*, 2, 564–565, November 1833

II. Account of the earthquake at Kathmandu. By A. Campbell, Esq. Assistant Surgeon, attached to the Residency.

On 26 Aug. last, at about 18:00, a smart shock of an earthquake was experienced throughout the valley and the neighbouring hills, westward in the valley of Nayakot and Duny Byas, eastward at Panouti, Baneppa, Dulkele and Pholam Chok, and southward at Chitlong, Chisagarhy, Etounda and Bissoulea. The shock was preceded by a rumbling sound to the eastward. The motion was undulatory, as of a large raft floating on the ocean, and the direction of the swell was from NE towards the SW. The shock lasted about 1 min. At 22:45 (Footnote: Not by chronometer, but by a good-doing clock which stopped during the great shock. Its pendulum vibrated north and south. [If the clock was set by the sun, the shock must have been 51 min earlier than Calcutta. – Ed.]) of the same day another shock of equal duration and of the same character occurred, and at 22:58, a third and most violent one commenced: at first it was gentle rocking motion of the earth, accompanied by a slight rumbling noise; soon, however, it increased to a fearful degree, the earth behaved as a ship at sea, the trees waved from their roots, and houses moved to and fro far from the perpendicular. Horses and other cattle, terrified, broke from their stalls, and it was difficult to walk without staggering as a landsman does on ship board. This shock lasted about 3 min in its fullest force. And the following is as correct an estimate as can be ascertained (without official documents) of the damage done by it to life and property throughout the great valley and neighbouring districts of Nipal. It is believed that the first two shocks were harmless.'

[Dr. Campbell's subsequent letters inform us that there have been frequent shocks of less violence since the above, many of which (on the 4 and 18 Oct. particularly) were felt at Calcutta, Monghyr, Chittagong, Allahabad and Jabalpur nearly simultaneously. On the 26th (Sept.), he writes, 'At 10:45 a sharp shock of the dangerous or undulating kind occurred. The embassy has returned from China, and I am informed that the great shock was not felt in Lassa, so that it would appear to have been confined to India within the Himalaya.' – Ed.]

Number of lives lost and buildings destroyed

Named location	Inferred location		Killed	Wounded	Houses
British Residency grounds			0	0	1
Kathmandu			60	38	400
Two pillars built by the minister, each upwards of 100 ft high, the large temple of Jagarnath, built by Ran Bahadur, after seven years of labour, and about a dozen temples were destroyed. The modern built garden houses of several members of the minister's family have been rendered untennantable, one of them, a handsome and ornamental edifice, came to the ground.					
South of the Capital					
Patan	Patan	27°39′, 85°19′	6	25	285
Sana gaon	Sanagau	27°38′, 85°22′	0	0	40
Harra Sidhi	Harisiddhi	27°37′, 85°21′	0	0	20
Teshu gaon			0	0	25
Selli gaon			0	0	16
Pagah	Panga	27°39′, 85°16′	0	0	24
Kuknah	Khokna	27°37.5,′ 85°18′	1	0	130
Baghmati	Bungamati	27°37′, 85°18′	0	0	80

contd...

A crack in the ground of 20 ft in length was observed at this village on the morning of the 27th; the entire number of houses in it was 206, more than a third of the whole were destroyed, and about 100 men [*more?*] were much damaged. The injury sustained here is probably greater than in any other part not to the east of Kathmandu.

Phurphing	Pharphing	27°36.5′, 85°16′	0	0	8
Chappa gaon	Capagau	27°36′, 85°19′	0	0	35
Peang	Pyanggau	27°35′, 85°19.5′	0	0	8
Taibu	Thaiba	27°36.5′, 85°21′	0	0	18
Bara gaon	Baregau	27°36′, 85°21.3′	0	0	35
Bali			0	0	3
Pahon			0	0	3
Sassanelly	Sisneri?	27°37′, 85°24′	0	0	2
Lubu	Lubhu	27°37′, 85°22.5′	0	0	25
Sana	Sanagau	27°38′, 85°22′	0	0	7
Hills about Sassanelly		27°35′, 85°24′	0	0	20

East of Kathmandu in the valley

Deo Patan	Deupatan				
Temple of Paspatnath escaped unhurt					
Handi gaon			0	0	20
Nag Desa	Nakdes	27°40.5′, 85°23′	4	0	20
Bareh	Bore	27°41′, 85°23′	5	0	20
Temi	Thimi	27°40′, 85°23′	0	0	150
Gou Karan	Gokarna	27°44′, 85°24′	0	0	8
Changu A fine old temple destroyed.			0	0	20
Sankhu	Sakhu	27°43′, 85°28′	20	5	45

A handsome temple of Mahadei, situated on a hill above Sankhu, was reduced nearly to ruins.

Bhat gaon	Baktipur	27°39′, 85°26′	200	104	2000

(Total number of houses in Bhat gaon is reckoned by Mr Hodgson at 4700, 3/4th of the town was said to be destroyed, 2000 being the average of many accounts, 6 or 8 fine temples were destroyed and a statue of Rajah Ranjit Mall, one of the Newar Princes of the Bhat gaon division of the valley.)

East of Kathmandu beyond the valley but nearby

Sangu	Sanga	27°38′, 85°29′	2	0	8
Baneppa	Banepá	27°37′, 85°32′	10	0	20
Nala gaon	Nala		6	0	11
Panouti	Panauti	27°34′, 85°31.5′	18	0	19
Dukele	Dhulikhel	27°37′, 85°35′	10	0	21
Phulam Chock (a fine temple destroyed here)		?Palanchock	60	0	300

NE of the valley

Dhulaka	Dolacka	86. 1°, 27.7′			
Listi gaon	Listikot	85.954°, 27.89′	0	0	0

Mundun Pahar, Kassa (Darm = Khasa, 85.95°27.00′), Kuti. In this direction the earthquake was much more severely felt. Kuti, a town on the Bhote frontier on the road to Lassa, is said to have been nearly destroyed, it contained 600 houses, only 50 of which remain. At Listi gaon (?Listikot), also on the Bhote frontier, a large portion of hill came down, and an iron bridge was destroyed. At Kan Sing Choke (?Choksum 86.00°28.07′), in the same direction, vulgar report says, that for 5 days before the earthquake took place, noises similar to the firing of cannons were heard as if underground: and in this neighbourhood the high road to Lhasa is said in many places to be blocked up by the fallen earth from the mountains. Shipa – the country residence of Colonel Runbir Sinh, 20 miles from Kathmandu on the Lhasa road, by the Kuti Pass, was seriously injured. Many small houses attached were destroyed and several lives lost.

West of Kathmandu

Swambunath	Swayanbhunath	27°42.3′, 85°17′	0	0	3

One small temple was destroyed, and the large one a little injured. The form of the large one must have preserved it. It is the chief buddhist temple in the valley, built in the fashion of that religion – an immense circular mound of brick work, surmounted by a four-sided spire or jweet.

Hal Chok	Dahacok	85°27′, 27.72°	0	0	3
Narod Devi			0	0	1

contd...

Changu Narayan			0	0	2
Goorkha Cantonment or Campoo (House of the Captain much injured)			0	0	4
Kirtipur	Kirtipur	27°40′, 85°16.5′	0	0	14
Contains 532 houses, and is built along the ridge and brow of a hill 3900 ft higher than the surrounding part of the valley. Its tenements are old and frail.					
Thankote	Thankot	27°41′, 85°12.5′	0	0	23
West of Kathmandu valley					
Duny byass neighbouring hills		10	0	40	
Tewanpur			0	0	10
Nayakoth	Nuwakot		0	0	3
North of Kathmandu					
Dharmtuli	Dharamtali	27°45.5′, 85°18′	0	0	2
Hukin gaon			0	0	1
Toka gaon	Tocka	27°45.5′, 85°20′	0	0	15
Burba Nil Kanth	Burhanikanth	27°46′, 85°21.5′	0	0	2
Chapaly	Capali	27°45.5′, 85°21.5′	0	0	7
Dharmpur	Dharmapur	27°44.5′, 85°20.5′	0	0	20
South of the valley					
Chitlong	Chitllong	27°40′, 85°11′	0	0	14
Chisapani	Chisapani	27°32′, 85°10′	0	0	0
The fort here was much injured: a large portion of breastwork facing the south had fallen, and the wall in many other places, although not fallen, was seriously injured.					
Mukwanpur	Mukwanpur	27°25′, 85°09′	0	0	0
The fort here has also suffered, but in much less degree than the fort at Chisapani.					
Total			414	172	4040

The above shows that the earthquake was much more severe to the north and east of the valley than here; and that even within the valley it was much more violent to the east of Kathmandu than at the capital itself, or other places to the west of it. The town of Bhat gaon is not more than 8 miles in a straight line from Kathmandu, and even there its violence must have greatly exceeded what it was at the latter place. To account for the immense disproportion in the loss of life and property at both the places, something may be allowed for the more frail state of the buildings at Bhat gaon; but this is not sufficient, and this circumstance must be considered as inexplicable as most others attending this fearful phenomenon. The brahmans of Nipal say (and it is believed with truth) that the occurrence of a more violent earthquake than this is recorded in their histories. It was about 600 years ago, and then the cities of Mangah, Patan and immumerable other towns were utterly destroyed and thousands of their inhabitants killed: the modern capital Kathmandu did not then exist.

Campbell, A., Further particulars of the earthquake in Nepal. *J. Asiatic Soc. Bengal*, VI, 636–639, December 1833

In pursuance of the attempt made before to note the destructive effects of the earthquake of 26 Aug. last throughout the valley of Nepal and its immediate neighbourhood, and with the hope of showing as correctly as my information will permit the probable seat or centre of this commotion, I beg to offer the following memoranda of other places at which the shock was experienced, as well as its comparative degree of intensity at each.

The means of estimating the violence of this phenomenon are, of course, most defective, if not wholly inadequate to the purpose: but in the absence of better data, the ascertained amount of damage done to the frail and perishable works of man may be received as an index of its intensity at one place, compared with that of another, and in conformity to this mode it would appear that the most extreme violence of the shock, as far its occurrence is yet known, was expended in the tract of country extending from this side of the great Himalaya range on the north to the course of the Ganges on the south, and from the Arun river (in the Nepal hills) on the east to the eastern branches of the Trisul Ganga on the west, comprising a space of about 200 miles from north to south and 150 miles from east to west. In this space, the valley of Nepal, though not geographically the centre point, is most assuredly the portion that has suffered the greatest violence of the calamity; and unless the inexplicable producing causes have been expended in the frequent and severe shocks that have to this day continued to recur, we may from our experience of the progress of the earthquake in other parts of the world, with reason, as we ought with resignation, look forward to further and more violent exhibitions of the same terrible nature.

In the notice of the earthquake by the secretary of the Asiatic Society, in his journal for August, he expressed a belief that the greatest intensity of the shock would be found to have occurred beyond the Himalaya, in the direction of Lassa; and judging from the direction from which the shock would be found to have proceeded and its intensity in the valley of Nepal, such was the probability, though other has turned out to be fact, and that upon good authority.

The recent return of an embassy from Nepal to the court of the celestial emperor has furnished authentic information on this subject, which otherwise might have been long wanting; and the hole tenor of it shows that the great Himalayan range itself and the country on *this*

side of it alone was the theatre of the earthquake's presence, and that it was not even in the slightest degree felt beyond a very short distance on the Tibetan side of those huge mountains.

The embassy was at Lassa on 26 Aug., when and where the shock was not experienced. At Digarchi, in the following month, it first received accounts of its occurrence from Nepal; to the inhabitants of that place the circumstance was known only from reports brought from this side of the mountains; along the road from Digarchi, the answer to all inquiries was the same. 'No earthquake on 26 Aug.', and not until its arrival at Tingri was it found that the shock had been felt. Tingri is a small Chinese post, immediately beyond the great Himalaya, and the first stage on the tableland (as it is called) of Tibet, going from hence to Lassa (by the Kuti or eastern pass from the Valley of Nepal). From Tingri to Kirung, a distance of 8 or 10 marches, the route is nearly due west, running along; and through the northern side of the Himalaya, and throughout this tract, though but thinly inhabited, authentic reports of the occurrence of the shock were received. By Kirung (the eastern pass from the valley into Bhote), the mission penetrated the great range, and at each stage (four in number through the pass) intelligence of the occurrence was communicated by the few individuals who inhabit that wild and sterile region. But such information was not required, as its effects were sufficiently manifest: in the village of Kirung itself, supposed to contain 400 houses, 60 were fairly demolished and many more seriously injured: two men had been killed under the ruins of their houses, and about a dozen injured.

From the exit of the pass to Kathmandu there are no towns along the route, and scarcely any villages; but at many places, insulated (?isolated) houses of the mountaineers had been thrown down, and the precipitous banks of hills and mountains had been hurled into the subjacent valleys.

This shows the extent of damage done towards the north, and enables us to fix the line of Tingri (Lat. 28°) as the northern limit of the earthquakes presence, and reports would show that of Jabalpur and Calcutta to have been the southern one. Rangpur defines the east (footnote from editor – Mr Walters informs me that it was also felt at Chittagong), and Delhi the west.

Northeast from Kathmandu as far as Dulka and Kuti, the violence would seem to have been greater than in the valley. West from Kathmandu it diminished at every step. At Gorkha, only two houses were destroyed; at Palpa, none; and at Doti, on the borders of Kumaon, the shock was felt, but by no means severely. It will strike everyone as remarkable that, while here the shock was more violent than elsewhere, its effects should not have been felt equally at as great a distance from hence to the north as to the south.

I subjoin an accurate register of the shocks which have occurred up to this date, given to me by Captain Robinson. Many of them have been severe, and throughout the whole course of these visitations, there have been two distinct varieties observed in the characters of the shocks, all those at the commencement were undulatory or swinging kind; the others wanted this swell, and were a violent up and down shaking with little vertical motion. The first may be called horizontal, the latter the vertical, variety. The former alone have been destructive to property, while the latter, from the greater noise by which they are accompanied, and the more rapid oscillations of the ground, are perhaps more terrifying.

Register of earthquakes at Kathmandu, from 26 Aug. to 26 Nov. 1833 inclusive

Aug. 26	17:55, 22:50, 22:58, duration 3 min, nine others, all horizontal
Aug. 27	04:53, 05:20, 05:26, horizontal
Aug. 28	07:25, 16:55, horizontal
Aug. 30	09:00 and 3 other shocks
Aug. 31	2 slight during the night
Sept. 1–11	10 slight shocks
Oct. 4	07:30, Smart vertical shock, 1 min duration, felt at Gorackpur and Allahabad
Oct. 18	13:15, severe, vertical, duration 1 min, felt slightly at Allahabad
Oct. 26	10:37, slight
Nov. 8	01:35, slight
Nov. 16	24:00, slight
Nov. 26	23:45, severe, vertical 1 min duration (occurring at the full moon)

1. Khattri, K. N. and Tyagi, A. K., *Tectonophysics*, 1983, **96**, 281–297.
2. Campbell, A., *J. Asiatic Soc. Bengal*, 1833, **2**, Misc. VI, 636–639.
3. Chitrakar, G. R. and Pandey, M. R., *Bull. Geol. Soc. Nepal*, 1986, **4**, 7–8.
4. Jackson, M. and Bilham, R., *Geophys. Res. Lett.*, 1994, **21**, 1169–1172.
5. Jackson, M. and Bilham, R., *J. Geophys. Res.*, 1994, **99**, 13897–13919.
6. Bürgmann, R., Bilham, R.., Gaur, V., Paul, J., Larson, K. and Jackson, M., *Eos Trans. Am. Geophys.*, 1994, **75**, 184–185.
7. Bilham, R., Bodin, P. and Jackson, M., *J. Nepal Geol. Soc.*, 1995, in press.
8. Sacks, I. S. and Linde, A. T., in *Earthquake Prediction: An International Review* (eds Simpson, D. W. and Richards, P. G.), Maurice Ewing Series, Amer. Geophys. Union, Washington, DC, 1981, vol. 4, pp. 617–628.
9. Beroza, G. C. and Jordan, T. H., *J. Geophys. Res.*, 1990, **95**, 2485–2510.
10. Khattri, K. N., *Tectonophysics*, 1987, **138**, 79–92.
11. Prinsep, H. *J. Asiatic Soc. Bengal*, 1833, **2**, Misc. IX, 438–439.
12. Campbell, A., *J. Asiatic Soc. Bengal*, 1833, **2**, Misc. II, 564–567.
13. Baird-Smith, R., *J. Asiatic Soc. Bengal*, 1843, **12**(2), New Series, 136, 1029–1056.
14. Baird-Smith, R., *J. Asiatic Soc. Bengal*, 1844, **156**, 964–983.
15. Mallet, R., Report of the Twenty-first Meeting of the British Association for the Advancement of Science, Ipswich, 1851, Second Report on the Facts of Earthquake Phenomena, 1852, pp. 272–320.
16. Mallet, R., Report of the Twenty-fourth Meeting of the British Association for the Advancement of Science, Liverpool, 1854, Catalogue of Recorded Earthquakes from 1606 BC to 1850 AD (continued from Report for 1853), 1855, pp. 2–326.
17. Oldham, T., Catalogue of Indian earthquakes, *Mem. Geol. Surv. India*, Geol. Surv. India, Calcutta, vol. 19, 1883, pp. 163–215.
18. Oldham, R. D., *Mem. Geol. Soc. India*, Geol. Surv. India, Calcutta, 1981, vol. 29, 379 pp.
19. Dunn, J. A., Auden, J. B., Gosh, A. M. N. and Roy, S. C., *Mem. Geol. Surv. India*, Survey of India, Calcutta, 1939 (reprinted 1981), vol. 73, 391 pp.
20. Bapat, A., Kulkarni, R. C. and Guha, S. K., *Catalogue of Earthquakes in India and Neighbourhood from Historical Period*

up to 1979, Indian Society Earthquake Tech., Roorkee, 1983, 211 pp.
21. Dunbar, P. K., Lockridge, P. A. and Whiteside, L. S., *Catalog of Significant Earthquakes 2150 BC to 1991 AD, Including Quantitative Casualties and Damage*, National Geophysical Data Center Report SE49, 1992, pp. 320.
22. Seeber, L. and Armbruster, J., in *Earthquake Prediction: An International Review* (eds Simpson, D. W. and Richards, P. G.), Maurice Ewing Series, Amer. Geophys. Union, Washington, DC, 1981, vol. 4, pp. 259–277.
23. Walker, J., Political Map of India, 1833, Military Secretary's Office, India House, Parbury and Allen, London, 1833.
24. India Gazette, *The India Gazette*, Calcutta, III, 3 Sept.–26 Oct. 1833.
25. Phillimore, R. H., *Historical Records of the Survey of India, 1800–1815*, Survey of India, Dehra Dun, 1950, Vol. II, pp. 477.
26. Sinha, K. K., in *Bihar–Nepal Earthquake 20 August 1988*, Geol. Surv. of India, Special Publication No. 31, 1993, pp. 104.
27. Pandey, M. R. and Molnar, P., *J. Nepal Geol. Soc.*, 1988, **5**, 22–44.
28. Rana, B. S. J. B., *Nepalko Bhukampa* (The Great earthquake of Nepal – *in Nepali*, 1935, 2nd edition, Kathmandu.
29. Dikshit, A. M. and Koirala, A., Report on the intensity mapping of Udaypur earthquake of 20 August 1988, HMG Ministry of Industry, Nepal, Dept. of Mines and Geology, Lainchaur, Kathmandu, May 1989.
30. Toppozada, T. R., *Bull. Seismol. Soc. Am.*, 1975, **65**, 1223–1228.
31. Hanks, T. C. and Johnston, A. C., *Bull. Seismol. Soc. Am.*, 1993, **82**, 1–23.
32. Johnston, A. C., Seismic moment assessment of earthquakes in stable continental regions, Part 1: Instrumental Seismicity, *Geophys. J. Internat.*, 1995 (in press).
33. Frankel, A., *Bull. Seismol. Soc. Am.*, 1994, **84**, 462–465.
34. Johnston, A. C., Seismic moment assessment of earthquakes in stable continental regions, Part 2: Historical Seismicity, *Geophys. J. Internat.*, 1995, (in press).
35. Johnston, A. C., *Geophys. J. Internat.*, in press.
36. Pandey, M. R. and Nicolas, M., Dept. of Mines and Geology, HMG Nepal, Report 2 March 1989, The Aftershocks sequence of the Indian–Nepalese earthquake of 20 August 1988.
37. Thakur, V. C. and Kumar, S., *Terra Nova*, 1994, **6**, 90–94.
38. Banerji, S. K., in *A Compilation of Papers on the Assam Earthquake of 15 August 1950* (ed. Ramachandra Rao, M. B.), Publication No. 1, Central Board of Geophysics, Govt. of India, 1953.
39. Chander, R., *Tectonophysics*, 1989, **170**, 115–123.
40. Molnar, P., 1990, *J. Himalayan Geol.*, 1990, **1**, 131–154.
41. Middlemiss, C. S., *Mem. Geol. Surv. India*, Geol. Surv. India, Calcutta, 1910, vol. 38, 409 pp.
42. Tandon, A. N., in *A Compilation of Papers on the Assam Earthquake of 15 August 1950* (ed. Ramachandra Rao, M. B.), Publication No. 1, Central Board of Geophysics, Govt. of India, 1953.
43. Poddar, M. C., in *A Compilation of Papers on the Assam Earthquake of 15 August 1950* (ed. Ramachandra Rao, M. B.), Publication No. 1, Central Board of Geophysics, Govt. of India, 1953.
44. Ray, in *A Compilation of Papers on the Assam Earthquake of 15 August 1950* (ed. Ramachandra Rao, M. B.), Publication No. 1, Central Board of Geophysics, Govt. of India, 1953.
45. Molnar, P., *J. Geol. Soc. India*, 1987, **30**, 13–27.
46. Gahalaut, V. K., Gupta, P. K., Chander, R. and Gaur, V. K., Minimum norm inversion of observed ground elevation change for slips on the causative fault during the 1905 Kangra Earthquake, Indian Acad. Sciences, in press.
47. Oldham, R. D., *Mem. Geol. Surv. India*, 1917, **XLII**, Part 2, 149–301.

ACKNOWLEDGEMENTS. I thank Arch Johnston for providing me with an advanced copy of his forthcoming paper, and Vinod Gaur for materials and insight on aspects of Himalayan seismicity. Librarians at the Norlin Library, Boulder, Colorado, and the India Office Library, London, were of great assistance in locating historic materials related to the 1833 earthquake. The research was funded by the National Science Foundation.

Correction

Is there dimorphism for style lengths in monoecious figs?

Prarthana Kathuria, K. N. Ganeshaiah, R. Uma Shaanker and R. Vasudeva
[*Curr. Sci.*, 1995, **68**, 1047–1049]

A recent change in the taxonomic position of the *Eupristina* (*Parapristina*) was brought to our notice by Dr U. C. Abdurahiman following the publication of our paper. The generic position of this pollinator wasp species has been recently shifted and hence on page 1047, Table 1, *Waterstoniella keralensis* (Abdurahiman and Priyadarsan) should read *Eupristina* (*Parapristina*) *keralensis* (Abdurahiman and Priyadarsan). Also, *F. tsjahela* should read *F. talboti*.

Erratum

The Great Bengal Cyclone of 1737 – An enquiry into the legend

A. K. Sen Sarma
[*Curr. Sci.*, 1995, **68**, 114–117]

The subtitle of the article, name and initials of the author were incorrectly printed in the Contents Page of the Journal. The correct subtitle, name and initials should read:

The Great Bengal Cyclone of 1737 – An enquiry into the legend

A. K. Sen Sarma
B-7/175, Kalyani
West Bengal 741 235, India

How fast is ultrafast chemistry?

Biman Bagchi

Solid State and Structural Chemistry Unit, Indian Institute of Science, Bangalore 560 012, India

The field of ultrafast chemistry has seen a string of remarkable discoveries in the recent years. In this article we briefly discuss some of the problems solved recently. The understanding that has emerged from these studies has important consequences not only in chemistry but also in diverse biological processes.

HOW fast is the response of a liquid to a given perturbation? The answer, of course, depends not only on the nature of the liquid but also on the nature of the perturbation itself. If the perturbation is slow, such as the shearing force on a liquid, then the response, the flow of the liquid, is also slow. And, if the perturbation is fast then the response is also fast. The question that a chemist has always asked is: How fast can this response be? What determines the limit? To a chemist, this question is not an idle one, but is profoundly related to the understanding of the effects of the solvent on the various chemical reactions[1–12] – we shall give examples later. And as the way we understand things now, the answer is also important to many biological reactions. Thus, we can formulate the question more precisely now: How fast is the response of chemically (and biologically) important solvents to various elementary chemical events that lead to much of the natural change around us?

The quest for the answer to the above question has occupied generations of chemists and finally the answer is beginning to take shape, thanks to some remarkable developments in this area in the last 3–5 years. For example, it was only very recently that the solvation dynamics of an ion in liquid water could be measured for the first time[11] – the dominant time constant is only about 50 fs (1 fs = 10^{-15} s)! No one had imagined that dynamics in water could be so fast. The importance of this particular experiment is that the dynamics was measured directly in the time domain. In fact, several such remarkable results have been obtained in the recent years. In the process many old ideas and notions have been found to be inadequate while new ones are still being tested. The objective of this article is to convey to the reader at least a part of the excitement of the field of ultrafast chemical dynamics.

There are many chemical reactions that occur at a great rate. It is only after the availability of ultrashort laser pulses and after the perfection of the associated detection techniques that study of these processes became possible. Figure 1 shows the evolution of time scales in the order they became available and a few examples of the dynamical processes which could be studied with the available time scales. It is important to realize that the underlying physical and chemical phenomena in different time scales are often profoundly different. In many cases the availability of shorter time scales meant more accurate measurement than was hitherto possible, often leading to new understanding. In some other cases completely new phenomena could be studied.

A key question in discussions on ultrafast processes is the limitation of the time scales posed by the Heisenberg uncertainty principle. When the time scales touch the femtosecond domain (the right bottom corner of Figure 1), then this certainly becomes a valid question. If we consider a relaxation that occurs with a time constant of 10 fs, then a quick calculation shows that the uncertainty principle imposes the limit that the minimum uncertainty in the frequency is 5×10^{13} Hz or, equivalently, 125 cm^{-1}. Thus, the uncertainty principle may make studies of vibrational relaxation irrelevant in some cases. However, the chemical processes that involve electronic relaxation will not be affected. There are,

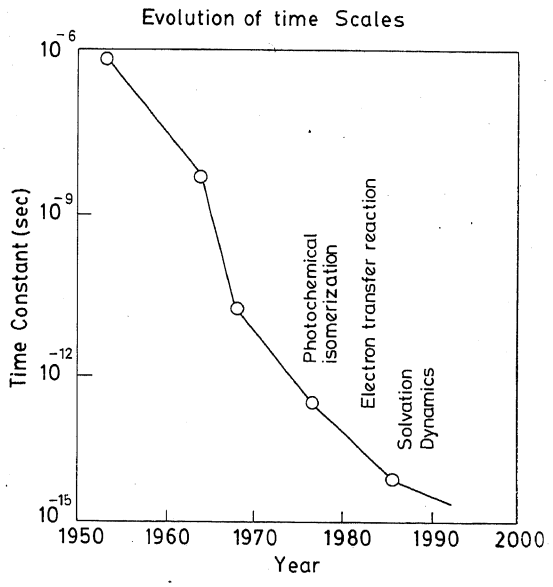


Figure 1. A schematic illustration of the evolution of time scales, where the experimentally accessible time constants of chemical processes are plotted against the years they became available. The three chemical relaxations discussed here, namely, isomerization dynamics, electron transfer reactions and solvation dynamics, are also shown in the same figure, approximately indicating the time when their modern studies became possible.

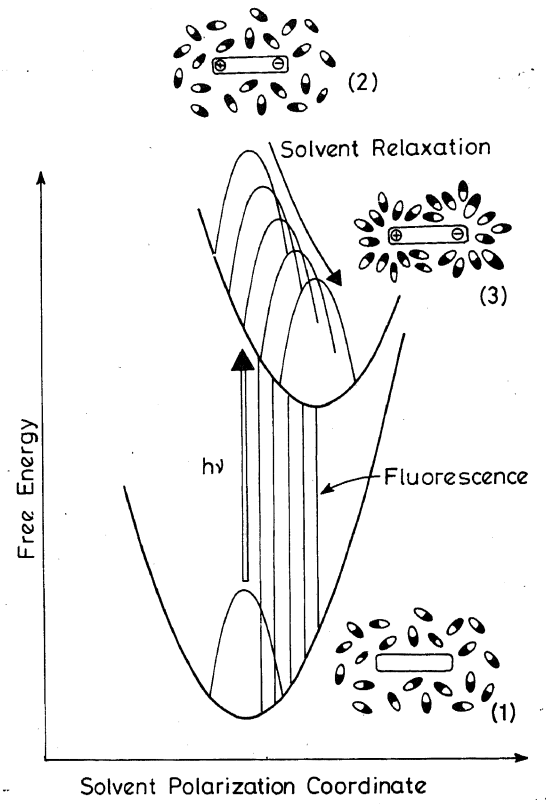


Figure 2. A schematic illustration of physical processes that accompany solvation dynamics and give rise to the time-dependent fluorescence Stokes shift (TDFSS). The state marked (1) is the unchanged, nonpolar state in equilibrium with the solvent. Optical excitation drives the system to the nonequilibrium state (2), where the solvent is still in the Franck–Condon state of the ground nonpolar molecule. Subsequent solvent reorganization leads to the stabilization (that is, solvation) of the polar molecule. As the system relaxes, the fluorescence from the solute shifts to the longer wavelength (that is, the spectrum gets red-shifted). The time dependence of this Stokes shift can be measured in experiments.

however, several other practical difficulties which make reliable study in the time scales below 20 fs difficult, at least at present. However, there may be interesting chemistry even in such short time scales. For example, the initial part of electron solvation dynamics after two-photon ionization of water molecules in liquid water occurs in times less than even 10 fs. During this time, the important process of the dephasing of the electronic wave function is virtually complete. It is not yet clear as to how to measure such processes.

In this review, we shall briefly touch upon a few of the chemical processes which have been studied intensely in the last few years. The examples are chosen from the author's own experience in the field, so they are not exhaustive. Also, the emphasis is entirely on the phenomena and their understanding – technical details (both experimental and theoretical) will be omitted. The list of references is not exhaustive. There are several reviews[1–8] and books[9, 10, 12] that have appeared in the last 5–6 years which contain the details and, in addition, provide useful starting point for the nonexpert. Boxes highlight the material expanding points briefly considered in the main text.

Solvation dynamics in dipolar liquids

The study of solvation of chemical and biological molecules is as old as chemistry and biology. Very few people, therefore, expected solvation dynamics to occupy the centre stage in ultrafast chemical dynamics. However, things turned out differently for various reasons. First, polar dye molecules serve as excellent probes in the study of solvation dynamics as these molecules have large polar stabilization energy in the excited state. This is because these dye molecules, upon excitation, often undergo a large increase in the dipole moment, or sometimes may even photoionize. Therefore, optical excitation prepares the molecule in the Frank–Condon state, which is of much higher energy than the minimum of the potential energy in the excited-state surface. This situation is shown in Figure 2. Subsequent to excitation, the solvent molecules rearrange and reorient to stabilize the new charge distribution in the excited state. The resultant change in energy is the solvation energy, which, as already mentioned, may be several electron volts. The second reason for the choice of dye molecules as suitable probes is that they often have easily detectable fluorescence with long life times, so one can easily study the time dependence of the Stokes shift of this fluorescence as the solvation energy of the dye molecules evolves. This time dependence is described by the solvation time correlation function, which is defined as[13]

$$S(t) = \frac{E_{\text{solv}}(t) - E_{\text{solv}}(\infty)}{E_{\text{solv}}(0) - E_{\text{solv}}(\infty)}, \qquad (1)$$

where $E_{\text{solv}}(t)$ is the solvation energy of the newly created ion at time t. This function is defined such that it

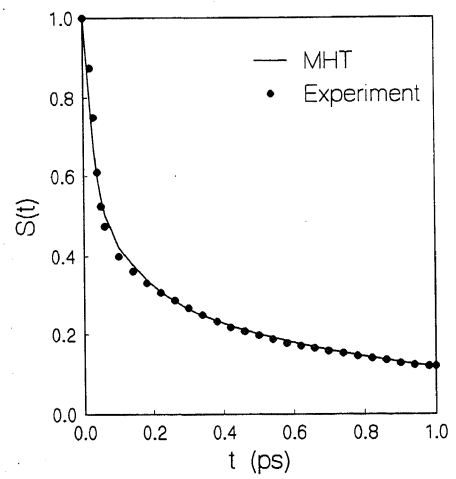
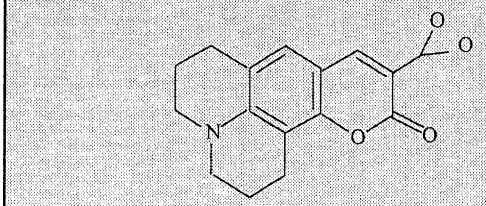


Figure 3. The solvation dynamics in liquid water. The black dots denote the experimental result[11], while the solid line is the theoretical prediction[24]. Note that the initial ultrafast part is over within 100 fs. For details, see the relevant references.

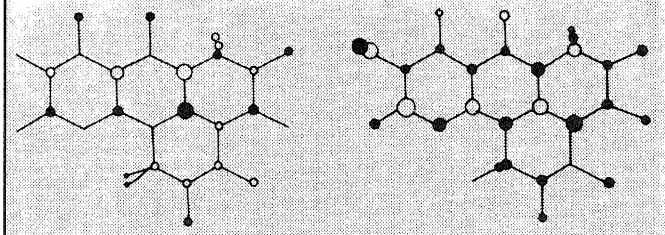
Box 1.

In the experiment done by Jimnez *et al.*[11] a Coumarin 343 dye (sodium salt, 10^{-4} M), in water, is excited with an ultrashort optical pulse. The evolution of the fluorescence spectrum of the Coumarin 343 anion (Figure B1.1 *a*) has been studied to follow the solvation dynamics of the anion. The atom-centred charges in the ground state and in the excited state, after excitation by optical pulse, are shown below in Figures B1.1 *b* and *c*, respectively.



$C343^-$

(*a*)



(*b*) (*c*)

The volume of each sphere is proportional to the charge on that atom and black denotes the negative charge. After excitation, the resultant fluorescence is resolved by a time-variable gate based on frequency-summing the fluorescence with a second ultrashort laser pulse (details given in the experimental section). As the solvent relaxation proceeds, there is a continuous shift of the emission spectrum towards the red end. The change in the peak frequency with time can be determined from these data. Thus, by observing the peak of the fluorescence spectrum at times t, zero and infinity, one can monitor the solvation dynamics of the anion and construct the solvation time correlation function $S(t)$ [vide equation (1)]. In the experiment a coherent Mira mode-locked Ti:sapphire laser has been used. The laser was operating at 850 nm. The second harmonic (425 nm), generated in a beta barium borate crystal, was used to excite Coumarin 343 (flowing in a quartz cell). The 850 nm light has been used to gate the fluorescence by mixing the fluorescence emission in another beta barium borate crystal. The cross-correlation of the two pulses (425 nm and 850 nm) was measured to be 100–110 fs (full width at half maximum). The intensity of the sum frequency has been measured as a function of the time delay between the two pulses. By angle-tuning the mixing crystal, emission decays were collected at certain intervals. The femtosecond fluorescence upconversion data obtained at a range of wavelenghts as a function of delay time has been plotted. These data can be used to construct a set of time-resolved emission spectra. The $S(t)$ can then be calculated from the set of spectra as discussed above.

decays from unity to zero as the solvation proceeds from $t = 0$ to $t = \infty$. The solvation dynamics is reflected in the red shift of the fluorescence spectrum and this experimental procedure is known as time-dependent fluorescence Stokes shift (TDFSS). One may expect that TDFSS will strongly depend on the nature of the solute probe, which can complicate the study of solvation dynamics. Fortunately, one finds that $S(t)$ is largely probe-independent, for reasons which are only partly understood, to be discussed later. The second reason for the great interest in solvation dynamics was the realization that it can significantly influence many chemical reactions, especially electron and proton transfers in polar liquids. Yet another reason is that the standard continuum model theories predicted that $S(t)$ would decay at a rate much faster than the rotational (and, of course, translational) relaxation. Thus, for water it predicted that the decay time would be about 250 fs (refs. 3, 13), which was startling even in the late eighties.

Over the last decade, $S(t)$ has been measured for a large number of polar molecules in different solvents. In the first phase (1986–1990), the experiments were done with time resolution only in the picosecond range. The solvation time correlation function was found to be non-exponential and differed from the predictions of the continuum models. During this period, an old comment of Onsager[14] became the topic of much discussion and we briefly discuss it here as it gave rise to interesting physics. In the Banff Conference on the 'Solvated Electron' in 1976, Onsager made the interesting comment that the polarization structure of an electron should form from outside in. That is, the molecules far from the electron should rearrange first, while the nearest-neighbour molecules relax the last. This suggestion came to be known as Onsager's 'inverse-snowball

Box 2.

The theory of solvation dynamics has undergone rapid development in the last few years to keep pace with the experiments. The most successful of the theories assume a linear response of the dipolar liquid. Under this assumption, the non-equilibrium response function $S(t)$ (defined by equation (1) of the text) is equal to the equilibrium solvation-energy–energy time correlation function, $C_{EE}(t)$, defined by

$$C_{EE}(t) = \frac{\langle E_{solv}(0)\, E_{solv}(t)\rangle}{\langle E_{solv}(0)\, E_{solv}(0)\rangle}. \qquad \text{(B2.1)}$$

This assumption leads to the simplification that one needs to consider only an equilibrium time correlation function, which is simpler to deal with. The microscopic calculation of $C_{EE}(t)$ is somewhat involved and the details are beyond the scope of this review. However, if one assumes that the main component of the solvation energy of an ion is the interaction between the bare electric field of the ion and the dipolar molecules of the liquid, then the final expression is simple and is given by[22-25]

$$C_{EE}(t) = \frac{\int_0^\infty dk \left(\frac{\sin kr_c}{kr_c}\right)^2 \mathcal{L}^{-1}\left[1 - \frac{1}{\varepsilon(k, z)}\right]}{\int_0^\infty dk \left(\frac{\sin kr_c}{kr_c}\right)^2 \left[1 - \frac{1}{\varepsilon(k)}\right]}, \qquad \text{(B2.2)}$$

where $\varepsilon(k, z)$ is the wavevector (k) and frequency (z) dependent dielectric function and $\varepsilon(k) = \varepsilon(k, z = 0)$ is the static dielectric function. r_c is a cut-off distance due to the short-range nature of the repulsive interactions, z is the (Laplace) frequency and $\mathcal{L}^{-1}$ denotes the Laplace inversion with respect to the frequency z. The standard dielectric constant is $\varepsilon_0 = \varepsilon(k = 0, z = 0)$. The above elegant expression has many satisfying features, one of them is that one recovers the well-known continuum model result simply by replacing $\varepsilon(k, z)$ by $\varepsilon(k = 0, z)$; the latter is the frequency-dependent dielectric function (z) measured in macroscopic experiments. The calculation of the wavevector- and frequency-dependent dielectric function forms the crux of the problem. This dielectric function $\varepsilon(k, z)$ can be shown to be given by the following expression[22]:

$$1 - \frac{1}{\varepsilon(k, z)} = \left(1 - \frac{1}{\varepsilon(k)}\right) \frac{1}{z + \Sigma(k, z)}, \qquad \text{(B2.3)}$$

where $\Sigma(k, z)$ is a generalized rate. This generalized rate derives contribution both from the rotational and the translational motions of the solvent molecules and also depends on the orientational pair correlation functions of the dipolar liquid. An important aspect of this rate is that in the absence of the translational contribution, it gives a smaller rate at intermediate wavevectors than at small values of k. This seems to agree with Onsager's 'inverse-snowball model' discussed in the text. However, when the translational component is efficient, then the complete reverse occurs, leading to the breakdown of the 'inverse-snowball model'. This theory has been applied to study solvation in many dipolar liquids. In Figure 3 the prediction of this theory has been compared with the experiments of Jimnez *et al.*[11] Note that the agreement is nearly perfect.

picture'. This also suggests that the solvation energy relaxation would be intrinsically nonexponential, as many, length-dependent, time scales are involved. Initial theoretical studies[15] supported this picture. However, it was argued later that the contribution of solvent translational modes may lead to the breakdown of the inverse-snowball picture[16]. More recently, it has been argued again that the Onsager picture may after all be correct for electron, but not for ions[17, 18]. The latest suggestion is that the electron being a very light particle will travel a long distance before it settles down in an existing trap. The dipolar molecules further from the electron can relax but the ones close to it must wait[18].

The most exciting experimental work on solvation dynamics has, however, taken place in the last 3 years only and the results will have far-reaching consequences. We shall refer to this period as the second phase. This started with the work on solvation of a dye, LDS-50, in acetonitrile reported from Fleming's laboratory at the University of Chicago[19]. This landmark experiment showed that solvation dynamics is biphasic with an ultrafast Gaussian component which decays with 80% of the total solvation energy within 200 fs. This is followed by a slow exponential-like decay with time constant of about 1 ps. This result was found to be in agreement with both computer simulation[20, 21] and theoretical[22] studies. However, the most interesting study, reported only last year, again from Fleming's group, is the solvation dynamics in water[11] (Box 1). As mentioned earlier, the results were remarkable. The solvation was found to be again biphasic. The initial Gaussian component decays with a time constant of less than 54 fs while the long time decay component can be described as a biexponential with time constants equal to 126 fs and 880 fs.

Figure 3 shows the solvation time correlation function in water. What is also interesting is the agreement of the long-time components with the earlier experimental results of Barbara and Jarzeba[3], who could not detect the ultrafast component because of limited time resolution. The solvation time correlation in methanol also contains an ultrafast Gaussian component with a time constant of about 70 fs. However, the relative weight of this part is somewhat less (about 30%) although still significant.

What is the origin of such ultrafast solvation in water, acetonitrile and methanol? These three liquids are quite different from each other and it is instructive to compare their dynamical features before we comment on the physical origin. Water, because of its small molecular weight and the extensive hydrogen bond network, has well-defined high-frequency vibrational modes. However, the single-particle orientation of water molecules is rather slow, with a correlation time of about 9 ps, due to the same hydrogen bonding. Acetonitrile, on the other hand, has fewer high-frequency vibrations, but orients very fast, with a correlation time of 0.3 ps. Methanol is again different. Because of chain-like hydrogen bond character, it has certain degree of high-frequency vibration, slow overall rotation but very fast rotation around the C–OH bond, which gives rise to a very fast polar response. All these details seem to be important in solvation dynamics[23].

We now turn to the explanation of the origin of the ultrafast component. Theoretical studies indicate[22–25] that all the natural fast dynamics of the system couple in different ways with the long-wavelength polarization fluctuation to give rise to the ultrafast component. This is because the main contribution to ion solvation energy comes from the long-wavelength part of the polarization, which is created in the liquid by the ion (Box 2). Now, the force constant for the longitudinal polarization fluctuation in a strongly polar dipolar liquid is large. Thus, the driving force to make ion solvation very fast exists in all strongly polar liquids. However, for solvation to be really fast, the liquid itself must be able to respond on the ultrafast time scale. This is where water and acetonitrile (and to smaller extent methanol and perhaps formamide) are exceptional. Water, due to its high-frequency vibrations and intermolecular vibrational modes, and acetonitrile, due to its fast rotational motion, can respond at a very high speed. The explanation given above is borne out by microscopic theory as shown in Figure 2.

Actually, even in these liquids, all is not ultrafast. Just as the rotation of water molecules is rather slow, so also the response to perturbations that vary on molecular length scales. These responses are also probed in many chemical processes, even in water. As discussed above, several rather unconnected (both microscopic and macroscopic) factors combine to give rise to the dominance of the ultrafast component in the solvation dynamics in water and acetonitrile.

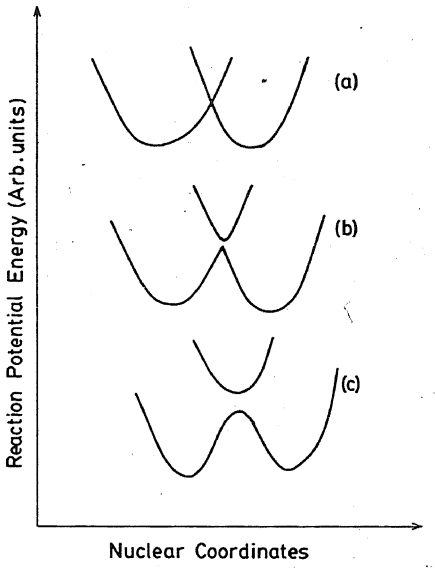


Figure 4. A schematic illustration of the nature of electron transfer reaction, based on the reaction potential energy surfaces. Depending on the strength of interaction between the two participating surfaces (corresponding to reactant and product states), the electron transfer can be classified as nonadiabatic, (case (a)), weakly adiabatic (case (b)) and adiabatic (case (c)).

Electron transfer reactions

Electron transfer reactions are ubiquitous in chemistry and biology. The initial studies were motivated by the desire to understand acid–base equilibrium and also many organic reactions involving electron transfer as a key step. With the advent of laser it is now possible to study many kinds of electron transfer reactions, especially those occurring in the excited state, for the first time. The reason is that many of these reactions were found to proceed at exceedingly fast speed, often with time constants in the subpicosecond range. Well-known examples are electron transfer in a photosynthetic reaction centre, fluorescence quenching in excited betaines and ion pair formation in various charge transfer complexes. We address below only the recent advances in understanding solvent effects in ultrafast electron transfer reactions in solution.

The key to the great success of the Marcus theory[26] lies in the choice of the reaction coordinate. The motion of the system on the reaction potential energy surface along this coordinate takes the system from the reactant surface minimum to the product surface minimum. In conventional chemical picture, this coordinate is usually some distance or angle and its nature is clear from the beginning. The situation is different for the electron transfer reaction. In this case, often no chemical bond is broken or formed. The major part of the activation energy comes from the interaction between the reaction system and the dipolar liquid and not from any intra-

molecular potential. Thus, the reaction coordinate is collective in origin. In the Marcus theory the reaction coordinate is essentially the solvation energy difference between the reactant and the product. However, in order to specify the nature of any electron transfer reaction we must know the nature of interaction between the reactant and the product surfaces. Interestingly, this interaction itself is determined by the nature of the Born–Oppenheimer surfaces as a function of the nuclear coordinates. Thus, in order to understand the electron transfer reaction one must consider the potential energy surface as a function of both the reaction coordinate and the nuclear coordinate. The reaction potential energy surface for several cases is shown in Figure 4. If the reactant and the product surfaces interact very weakly, then the reaction can be considered as a transfer of the electron from the former to the latter at the crossing point and the reaction is termed 'nonadiabatic'. On the other hand, if the two surfaces interact strongly near the crossing, then the electron transfer occurs on one surface (which is given by a combination of the two parent surfaces according to the rules of quantum mechanics) and the reaction is termed 'adiabatic'. The experimental situation often lies between these two broad classifications and is termed loosely as 'weakly adiabatic' or even 'weakly nonadiabatic'. This intermediate situation is shown in Figure 4 *b*.

Recent studies[27–33] have critically examined the dynamic effects of solvent on various electron transfer reactions (Box 3). One is particularly interested to know what effects the newly discovered ultrafast solvation could have on electron transfer reactions. The answer depends on the 'adiabaticity' of the reaction. For nonadiabatic reactions even the ultrafast polar response was found to be not fast enough to affect electron transfer seriously. The situation is rather different for weakly adiabatic electron transfer. Here the first and foremost effect is that the rate of barrier crossing increases by almost an order of magnitude over the old estimates that ignored the ultrafast solvation effects[31–33]. This is true not only for water, but also for the other two ultrafast solvents, acetonitrile and methanol. For acetonitrile some additional interesting features appear. For this solvent the reactive friction (which arises from coupling between the electronic charge and the solvent) becomes so small that the electron transfer is controlled not by the rate of barrier crossing but by the rate of energy diffusion to the barrier from the reactant well. Electron transfers in water and methanol are predicted to be in the normal regime. In fact, it is rather ironic that ultrafast solvation in water implies that this solvent will be able to respond very fast to chemical changes during a reaction which, in turn, will not show any dynamic solvent effects[32, 33].

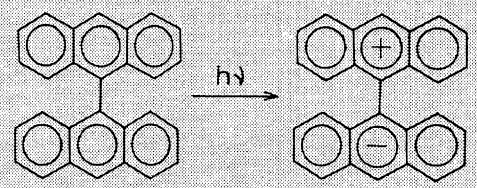
The Marcus theory predicts a parabolic dependence of reaction rate on the free energy of the electron transfer reaction (Box 4). Although this dependence was predicted in 1956, its verification had to wait till the 1984 landmark experiment of Miller *et al.*[34]. These authors measured the rate of electron transfer across a bridge of organic groups of varying length. Subsequently, several confirmations of the parabolic dependence have been reported[35]. However, there are also several recent studies of ultrafast electron transfer which found interesting breakdown of the Marcus parabolic dependence[36–40], due to various reasons. One such example is shown in Figure 5, where the rate of charge recombination (CR) of newly formed contact ion pair (CIP) is plotted against the free energy of reaction. The rate increases in the normal region with decrease in the free-energy gap. Murata and Tachiya[40] have recently provided an elegant explanation of this non-Marcus energy gap dependence by invoking the interplay between relaxation and electron transfer. A more detailed theoretical work by our group[41, 42] has shown that the above interplay can give rise to highly nonexponential decay of the reactant population.

Box 3.

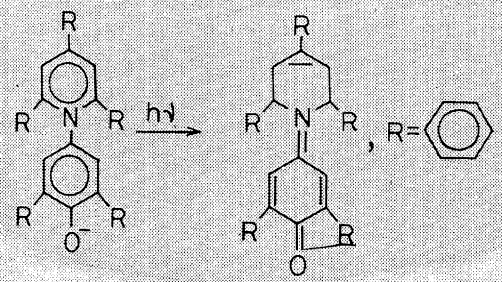
A very large number of experimental studies has been carried out in the recent years to understand various factors that influence electron transfer reactions. Three extensively studied examples are shown below:

(a) $Mn^{2+} + Mn^{3+} \Leftrightarrows Mn^{3+} + Mn^{2+}$

(b)



(c)



where (a), (b) and (c) correspond to the normal, zero-barrier and the inverted regime of electron transfer, respectively. Many different experimental techniques are being used routinely to study the dynamics of electron transfer. These include the rate of disappearance of fluorescence from the reactant well (due to electron transfer), femtosecond excited-state absorption and also the ground state recovery.

Box 4.

In the Marcus theory, the reaction coordinate (X) is a collective quantity and is defined by the following expression:

$$X = \int d\mathbf{r}\, \Delta \mathbf{D}(\mathbf{r}) \cdot \mathbf{P}(\mathbf{r}), \qquad \text{(B4.1)}$$

where $\Delta \mathbf{D}(\mathbf{r})$ is the change in the displacement vector due to the charge redistribution owing to the electron transfer and $\mathbf{P}(\mathbf{r})$ is the polarization field. Note that here the reaction coordinate itself is an energy. The reaction potential energy has minima corresponding to the reactant and the product states. In the Marcus theory, electron transfer is considered to be mediated by solvent polarization fluctuations and distinction is made between the nonadiabatic and adiabatic electron transfers. For the former case, the rate is given by the following expressions:

$$k = (V^2/\hbar)\,(4\pi\lambda / k_B T)^{1/2} \exp(-\beta E_a), \qquad \text{(B4.2)}$$

where V is the nonadiabatic coupling between the two surfaces (determined by the nuclear coordinates). The activation energy E_a has the following well-known form:

$$E_a = (\Delta G^0 + \lambda)^2 / 4\lambda, \qquad \text{(B4.3)}$$

where ΔG^0 is the free energy gap and λ is the solvent reorganization energy. In Figure B4.1 we show the potential energy surface to explain the notation.

In the adiabatic limit, the rate expression is different and is given by

$$k = \frac{\omega_R}{2\pi} \exp(-\beta E_a), \qquad \text{(B4.4)}$$

where ω_R is the harmonic frequency of the reactant well.

The above expressions are essentially transition state expressions. The important thing to note is the parabolic trend of the rate predicted by equation (B4.3) as a function of the free-energy gap ΔG^0.

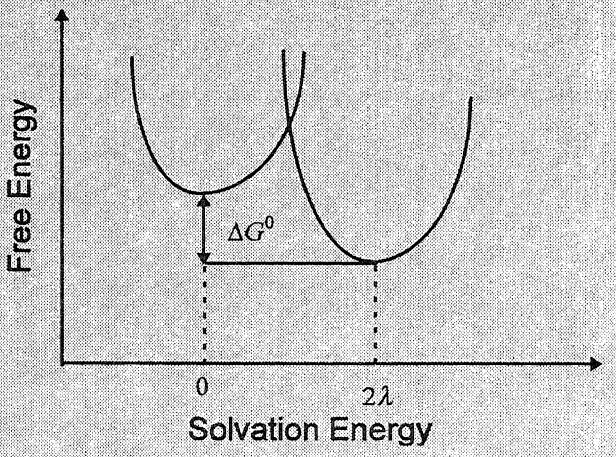


Figure B4.1. A schematic representation of the reaction free-energy surface, according to the Marcus theory. ΔG^0 is the free-energy gap between the reactant and the product and λ is the solvent reorganization energy. Note that the reaction coordinate (the abscissa) has been set to solvation energy.

Since the reaction coordinate is the fluctuating energy gap due to solvation, it is obvious that solvation dynamics will play an important role in determining the dynamics of electron transfer. Zusman was the first to quantify this dependence by deriving the following expression for an adiabatic reaction:

$$k = \frac{1}{\tau_L}(16\pi k_B T/\lambda)^{-1/2} \exp(-\beta E_a), \qquad \text{(B4.5)}$$

where τ_L, the longitudinal relaxation time, is a measure of the solvation time of a newly created ion[1–10].

When the solvent polar response is ultrafast, like in water and acetonitrile, then the above expression for solvent effects undergoes significant modification. The theoretical work that accounts for ultrafast solvation is available in ref. 32. The main results are described in the text.

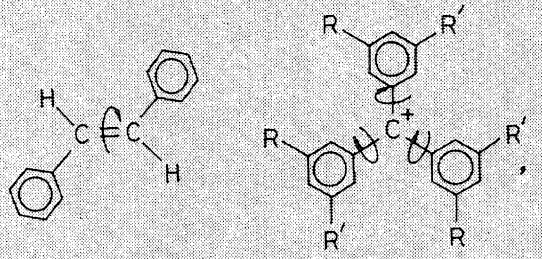
Another reason for the non-Marcus energy dependence is the involvement of high-frequency modes of the product state, which often serve as the accepting modes to accelerate electron transfer[37–39]. This mechanism is particularly useful when the electron transfer is in the Marcus inverted regime and results in the absence of the decrease of the rate as the free-energy change between the reactant and the product is made larger (that is, G is negative and large). This mechanism of the non-Marcus energy gap dependence seems to be responsible for the large electron transfer rate observed recently in betaines[38].

Isomerization dynamics

Isomerization reactions are very important chemical reactions. There are several reasons why these reactions often proceed at a very high speed. First, no chemical bond is broken or formed. The barrier is often rather small, only a few kcal/mol[43]. Moreover, there are several important reactions which proceed in the absence of any activation barrier. Some examples are the *cis* → *trans* isomerization of excited stilbene and also the isomerization of triphenyl methane dyes, such as malachite green[43]. These reactions are naturally very

Box 5.

There are two classes of molecules which are particularly useful in the study of isomerization dynamics. These are stilbenes and the triphenyl methane dyes. The structures are shown below:



R, R′ = alkyl group

The reactive motion, in each case, is a twist of a group around a body-fixed axis, as shown by the arrows. The isomerization reaction is a high-barrier one for *trans* → *cis* isomerization in stilbene, but nearly barrierless for *cis* → *trans* for the same system. It is also barrierless in a large number of TPM dyes, like crystal violet and Malachite green.

Experimentally, these molecules are often studied by exciting them to a higher electronic state. The dynamics of isomerization is then studied by monitoring the decay of fluorescence from these dyes. Since isomerization leads to the formation of a new state, the fluorescence from the reactant gets quenched as the reaction proceeds. Thus, the time dependence of the decay of fluorescence intensity serves as an excellent probe of isomerization dynamics. The experiments have usually concentrated on the temperature and viscosity dependence – the latter often shows very interesting fractional viscosity dependence. However, the latter problem, despite considerable effort over the last two decades, has remained largely ill-understood. This is because the potential energy surfaces of isomerization reactions appear to be rather sensitive to their environment. For example, the activation energies change significantly when the solvent is changed. Even the change of temperature affects the activation energy because the properties of the solvent often change significantly with temperature. This is also true, perhaps to a lesser extent, for pressure change. Thus, one is left with very limited scope for changing the viscosity of the solvent without affecting the reaction potential energy surface. Thus, computer simulations appear to be the ideal method to bridge the gap between theory and experiments.

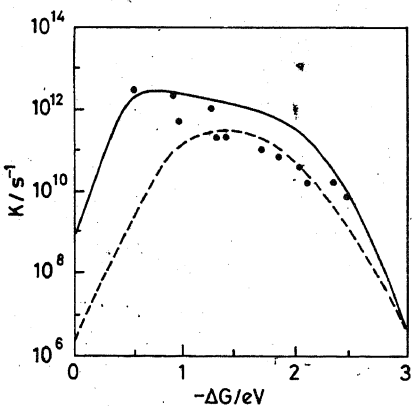


Figure 5. Breakdown of the Marcus parabolic dependence on the free-energy gap for charge recombination in newly formed contact ion pair. The black circles denote the experimental results of Asahi and Mataga[36] while the solid line is the theoretical prediction of Murata and Tachiya[40]. The Marcus parabola is also shown by the dashed line.

fast. There are some cases also where the barrier is significant. In other words, the isomerization reactions span a diverse range in kinetic behaviour. Because of the simplicity of these reactions, they serve as the testing grounds for the theoretical models of solvent effects. We shall discuss here two important and recent theoretical developments which were initially motivated by experimental results on ultrafast isomerization reactions in solution but led to far-reaching consequences (Box 5).

Before 1980, the effects of solvent viscosity on isomerization reaction were explained by using the well-known Kramers' theory[44]. This theory assumes that a chemical reaction can be viewed as the passage of a Brownian particle over an activation barrier. For large solvent viscosities, this theory predicts that the rate is inversely proportional to the viscosity. This is a straightforward consequence of modelling the reaction as a Brownian motion. Initial experiments seemed to have supported this. Detailed experimental studies could be undertaken only after picosecond laser became available in the early eighties. The new experimental results clearly indicated that for fast isomerization reactions (with lifetimes in the picosecond and even in the subpicosecond range) in solution, Kramers' theory was inadequate to explain the viscosity dependence – the observed viscosity dependence was much weaker than the inverse dependence predicted by Kramers' theory[45–47].

Box 6.

The recent experimental results from the study of isomerization dynamics have played a key role in the development of modern chemical reaction theory. Not only did these results motivate newer studies, but they also serve as the testing ground for the older theories. Here we discuss one particular class of theories which have been quite successful in explaining the viscosity dependence of the rates. Prior to 1980, the only theoretical framework available in this area was that due to Kramers, who modelled a chemical reaction as the motion of a solute in a one-dimensional double well potential. The reaction occurs when the solute escapes from one well to the other and in the process needs to go over the potential barrier, which is modelled as an inverted parabola of frequency ω_b. Kramers assumed that the equation of motion was an ordinary Fokker–Planck equation. This second-order partial differential equation could be solved for the steady-state solution, which gives the following expression for the rate:

$$k = \frac{\omega_R}{2\pi\omega_b}\left[\left(\frac{\zeta^2}{4}+\omega_b^2\right)^{1/2}-\frac{\zeta}{2}\right]\exp(-\beta E_a), \quad \text{(B6.1)}$$

where E_a is the activation energy, ω_R and ω_b are the harmonic frequencies of the reactant well and of the barrier top, respectively. ζ is the friction on the reactant motion. For isomerization reaction, this friction may be proportional to the viscosity of the solvent. Kramers' expression predicts that at high viscosities (that is, larger than 10 cP), the rate should be proportional to viscosity. However, experiments indicate that for many reactions with high barrier, the rates show a much weaker (such as, fractional) viscosity dependence. In 1980, Grote and Hynes[49] presented a generalization of Kramers' theory. These authors pointed out that Kramers' theory may be inadequate because it neglects the memory of motion along the reaction coordinate. The generalized rate is obtained by solving the following relation self-consistently:

$$\lambda_r = \frac{\omega_b^2}{\lambda_r + \hat{\zeta}(\lambda_r)}, \quad \text{(B6.2)}$$

where the factor λ_r gives the required rate by $k = \lambda_r k^{TST}$, k^{TST} being the transition state rate. In the above equation, $\hat{\zeta}(x)$ is the frequency-dependent friction, which gives the viscoelastic response of the solvent. As explained in the text, a reaction with a sharp barrier can probe only the high-frequency component of friction, which can be vastly different and much smaller than the zero-frequency friction that enters in Kramers' theory. This can explain the weak viscosity dependence of rate in many isomerization reactions. However, as described in Box 5, experimental difficulties have not yet allowed a full, convincing comparison between theoretical predictions and the experiments.

This started a flurry of theoretical activity[48–51]. Interestingly, the work that proved to be the most useful in this stage was published only a few years ago by Grote and Hynes[49]. The understanding that has emerged is the following. In *cis* ↔ *trans* isomerization reactions, the two potential energy minima are often separated by a dihedral angle lying between 90° and 180°. This gives rise to a sharp barrier when the activation energy barrier is somewhat larger than 10 kcal/mol. A sharp activation barrier implies that the reactant spends very little time in the barrier region and, therefore, probes only the high-frequency part of the solvent frictional response. The relation between the barrier height and the barrier curvature has been illustrated in Figure 6. This is to be contrasted with Kramers' theory, which assumes a slow diffusive motion along the reaction coordinate across the barrier (Box 6). For a sharp barrier only the former picture is valid. Since at large viscosities the high-frequency frictional response of the liquid gets decoupled from the macroscopic (that is, zero frequency) viscosity, the reaction rate also gets decoupled from viscosity and this shows up as a weak dependence of the reaction rate on the solvent viscosity. This rather elegant explanation is also known as the non-Markovian effect in barrier crossing.

Another class of isomerization reaction has drawn attention in the recent years[43, 52–54]. This is the zero-barrier reaction limit. As already mentioned, several important photochemical reactions fall in this limit. Consider that a molecule is optically excited and that in the excited state there is a photochemical funnel at the minimum of the excited-state surface. The motion that takes the system from its initial state to the final state is the reactive motion. The situation is depicted in Figure 7. This reactive motion can be the rotation of a bulky, molecular group (as in the case of TPM dyes). In this case the rate of reaction can be very high and dependent on the solvent viscosity. For *cis* → *trans* isomerization of stilbene in hexane, the isomerization rate seems to be as high as 10^{13} s^{-1} (refs. 53, 54). This is one of the fastest rates

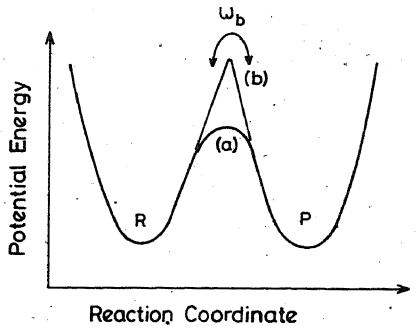


Figure 6. A schematic illustration of the dependence of the curvature of the activation barrier of the isomerization reaction on the barrier height. As the barrier height increases, the barrier frequency (ω_b) also increases, leading to the decoupling of the rate from the solvent viscosity.

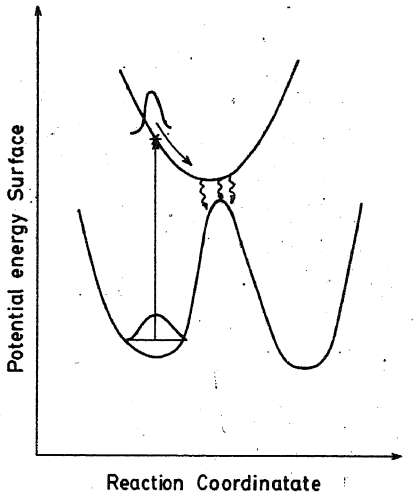


Figure 7. A schematic illustration of a zero-barrier chemical reaction. Initial excitation places the molecule in a nonequilibrium position on the excited-state potential energy surface. Subsequent relaxation brings the molecule to the potential minimum, where a photochemical funnel (or a sink) is present, This funnel gives rise to the decay of the excited state. The reactive motion is the relaxation from the initial excitation position to the configuration where the sink is located.

measured in the condensed phase. The decay of population on the excited state can often be measured by studying the disappearance of fluorescence from the excited state or by the recovery of the ground state population. An important characteristic of this type of reaction is that the time dependence of the excited-state population is highly nonexponential. Another important hallmark of barrierless reactions is the fractional viscosity dependence of the rate. The reason for the fractional viscosity dependence here is rather different from that in the high-barrier reaction and is often the multidimensional nature of the reaction potential energy surface; the second dimension may be a solvent mode or an intramolecular vibrational mode.

The theoretical studies of barrierless reactions are mostly based on a stochastic approach, where a Smoluchowski or a Fokker–Planck equation is used to describe the motion on the reaction potential energy surface with a coordinate-dependent sink term to account for the decay from the photochemical funnel. This approach[43] seems to have been fairly successful in describing the experimental results.

Ultrafast chemistry in biological processes

It should be pointed out that it is a fairly recent realization that the barrierless chemical reactions are quite common not only in chemistry but also in biological reactions. The most famous examples are the electron transfer in the reaction centre in photosynthesis[55] and the isomerization reaction in rhodpsin, which triggers the proton pump in the retinal membrane. These discoveries have led to very interesting chemistry. In what follows, we briefly discuss a very recent theoretical work by Chandler *et al.*[56] that led to valuable insight into the electron transfer mechanism in the photosynthetic reaction centre. A schematic representation of the relevant part of the reaction centre is shown in Figure 8. It is known that after the photoexcitation, the electron transfer takes place from 1 to 3 along the right-hand side of the figure. The theoretical work mentioned above has suggested that the reason for this selective transfer is that this reaction is barrierless while the $1 \rightarrow 2$ transfer involves a sizeable activation barrier[56]. A schematic drawing of this situation is shown in Figure 9. Another important example is the isomerization of bacteriorhodopsin in the retinal membrane, halobrium halobacterium. This is also a barrierless reaction with a time constant of 0.5 ps (refs. 57, 58). This reaction has recently raised a lot of interest. Yet another example of much current interest is the recombination of oxygen and carbon monoxide with the iron in heme in myoglobin[59, 60]. Flash photolysis studies have shown that this reaction follows a nonexponential kinetics. Interesting theoretical work has been done on this problem. Much experimental, computer simulation and theoretical studies have been done on the dynamics of proteins, which seem to have relaxation over a range of time scales. As the technique of ultrafast spectroscopy becomes more and more refined, we can look forward to many more fascinating applications of ultrafast chemistry to biological systems.

In the past many experts have raised the following question: What could be the relevance, if any at all, of the ultrafast chemistry to biology when the life processes occur on much slower time scales? This is certainly a valid question. With the recent discoveries of ultrafast biological processes (such as electron transfer in a photosynthetic reaction centre), this question needs to be reformulated as follows: Why does Nature use so

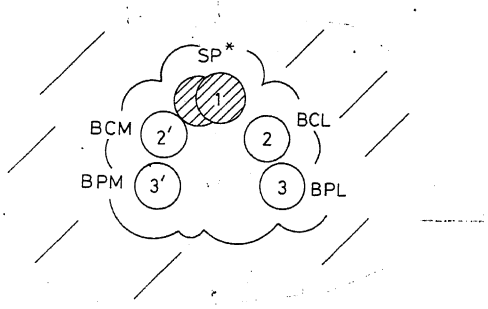


Figure 8. A schematic depiction of the essentials of the primary electron transfer event in a bacteria photosynthetic reaction centre. SP* denotes the photoexcited special pair. The electron, in principle, can be transferred either to the L branch, forming the charge-separated states SP^+ BCL^- or SP^+ BPL^-, or it can go to the M branch, forming similar states. The experimentally observed case is SP^+ BPL^-.

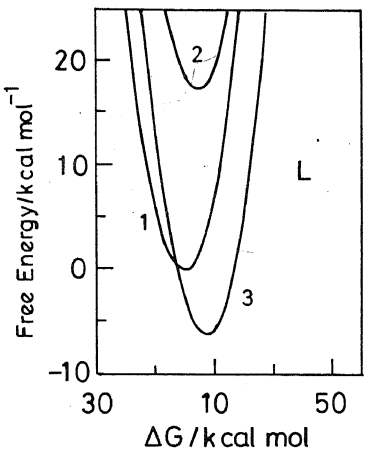


Figure 9. A schematic illustration of the explanation given by Chandler *et al.*[56] for the formation of SP^+ BPL^- (see Figure 8). Here the state SP* is the electronic state 1, SP^+ BCL^- is the electronic state 2 and SP^+ BCL^- is the electronic state 3. This figure shows the calculated Marcus free energy for the diabatic surfaces for all the three states. The electron transfer occurs from 1 to 3 because it is barrierless and energetically favourable. The transfer to M branch ($1 \rightarrow 3$) involves a sizeable activation barrier.

many different time scales in biological functioning? Qualitatively, the answer to this apparently profound question is simple enough: the dynamical processes at different time scales are hierarchially coupled. An extreme example of this coupling is provided by diffusion in the supercooled liquid. Here the rapid rattling motion of a molecule in the cage formed by its neighbours is ultimately responsible for the rare jumps or hops that give rise to the diffusive mass transport. The time scales of these two motions differ by many orders of magnitude, but the slow motion would be impossible without the former. Perhaps the reason that Nature uses so many different time scales, from ultrafast to ultraslow – often separated by several orders of magnitude – is to minimize the scope of making mistakes. This is probably developed by natural selection over many years.

Conclusion

With the advancement of ultrafast spectroscopy in the last decade, many important discoveries have been made which have helped in solving some long-standing puzzles in chemistry and biology. For example, one was puzzled by the absence of dynamic solvent effects on various charge transfer reactions in water. The paradox was partly due to the known result that the dielectric relaxation in water is about 9 ps, which is quite long. Only last year was it discovered that solvation dynamics in water is more than two orders of magnitude faster. This explains why we do not see the above-mentioned solvent dynamic effects in water. Similarly, the recent understanding of electron transfer reactions, especially the involvement of the high-frequency vibrational modes, has also emerged only in the last five years and again the impetus came from the discovery of novel aspects by using ultrafast spectroscopy. With the availability of femtosecond lasers in many laboratories around the world, we can look forward to many more such interesting discoveries in the near future.

1. Fleming, G. R. and Wolynes, P. G., *Phys. Today*, 1990, **43**, 36.
2. Bagchi, B., *Annu. Rev. Phys. Chem.*, 1989, **1989**, 115.
3. Barbara, P. F. and Jarzeba, W., *Adv. Photochem.*, 1990, **15**, 1.
4. Maroncelli, M., McInnis, J. and Fleming, G. R., *Science*, 1990, **243**, 1990.
5. Bagchi, B. and Chandra, A., *Adv. Chemys.*, 1991, **40**, 1.
6. Maroncelli, M., *J. Mol. Liq.*, 1993, **57**, 1.
7. Y. Gauduel, *J. Mol. Liq.*, 1995, **63**, 1.
8. Rossky, P. J. and Simon, J. D., *Nature*, 1994, **370**, 263.
9. Gauduel, Y. and Rossky, P. J. (eds), *Ultrafast Reaction Dynamics and Solvent Effects*, AIP Press, New York, 1994.
10. Simon, J. D. (ed.), *Ultrafast Dynamics of Chemical Systems*, Academic, Kluwer, 1994.
11. Jimnez, R., Fleming, G. R., Kumar, P. V. and Maroncelli, M., *Nature*, 1994, **369**, 471.
12. Fleming, G. R., *Chemical Applications of Ultrafast Spectroscopy*, Oxford University Press, Oxford, 1986.
13. Bagchi, B., Fleming, G. R. and Oxtoby, D. W., *Chem. Phys.*, 1986, **86**, 257.
14. Onsager, L., *Can. J. Chem.*, 1977, **55**, 1819.
15. Calef, D. F. and Wolynes, P. G., *J. Chem. Phys.*, 1983, **78**, 4145.
16. Chandra, A. and Bagchi, B., *J. Chem. Phys.*, 1989, **91**, 2594.
17. Tachiya, M., *Chem. Phys. Lett.*, 1993, **203**, 164.
18. Roy, S. and Bagchi, B., *Chem. Phys.*, 1994, **183**, 207.
19. Rosenthal, S. J., Xie, X., Du, M. and Fleming, G. R., *J. Chem. Phys.*, 1991, **95**, 4715.
20. Maroncelli, M., *J. Chem. Phys.*, 1991, **94**, 2084.
21. Neria, N. and Nitzan, A., *J. Chem. Phys.*, 1992, **96**, 5433.
22. Roy, S. and Bagchi, B., *J. Chem. Phys.*, 1993, **99**, 1310, 9938.
23. Roy, S. and Bagchi, B., *Adv. Chem. Phys.*, 1995, under preparation.
24. Nandi, N., Roy, S. and Bagchi, B., *J. Chem. Phys.*, 1995, **102**, 1390.
25. Roy, S., Komath, S. S. and Bagchi, B., *J. Chem. Phys.*, 1993, **99**, 3139.
26. Marcus, R., *Annu. Rev. Phys. Chem.*, 1964, **15**, 153.
27. Zusman, L. D., *Chem. Phys.*, 1980, **49**, 295.
28. Hynes, J. T., *J. Chem. Phys.*, 1986, **90**, 3701.
29. Weaver, M., *Chem. Rev.*, 1992, **92**, 463.
30. Barbara, P. F., *et al.*, *J. Chem. Phys.*, 1992, **96**, 3728, 7859.

31. Smith, B. B., Staib, A., Hynes, J. T., *Chem. Phys.*, 1993, **176**, 521.
32. Roy, S. and Bagchi, B., *J. Phys. Chem.*, 1994, **98**, 9207.
33. Roy, S. and Bagchi, B., *J. Chem. Phys.*, 1995, in press.
34. Miller, J. R., Calcaterra, L. T., Closs, G. L., *J. Am. Chem. Soc.*, 1984, **106**, 3047.
35. Mataga, N., Asahi, T., Kanda, Y., Okada, T. and Kakitani, T., *Chem. Phys.*, 1988, **127**, 249.
36. Asahi, T. and Mataga, N., *J. Phys. Chem.*, 1991, **95**, 1956.
37. Bixon, M. and Jortner, J., *Chem. Phys.*, 1993, **176**, 467.
38. Tominaga, K., Walker, G. C., Jarzeba, W. and Barbara, P. F., *J. Phys. Chem.*, 1991, **95**, 10475.
39. Nagasawa, Y., Yartsev, A. P., Tominaga, K., Johnson, A. E. and Yoshihara, K., *J. Am. Chem. Soc.*, 1993, **115**, 7922; Kandori, H., Kemnitz, K. and Yoshihara, K., *J. Phys. Chem.*, 1992, **96**, 8042.
40. Murata, S. and Tachiya, M., *J. Am. Chem. Soc.*, 1994, **116**, 2434.
41. Gayathri, N. and Bagchi, B., *J. Mol. Struc. (Theor. Chem.)*, 1995, in press.
42. Gayathri, N. and Bagchi, B., *J. Chem. Phys.*, submitted.
43. Bagchi, B., *Int. Rev. Phys. Chem.*, 1987, **6**, 1.
44. Kramers, H., *Physica*, 1940, **7**, 284.
45. Velsko, S. P., Waldeck, D. H. and Fleming, G. R., *J. Chem. Phys.*, 1983, **78**, 249.
46. Rothenbergere, G., Negus, D. K. and Hochstrasser, R. M., *J. Chem. Phys.*, 1983, **79**, 5360.
47. Flom, S. R., Brearley, A. M., Kahlow, M. A., Nagarajan, V. and Brabara, S. R., *J. Chem. Phys.*, 1985, **83**, 1993.
48. Nitzan, A., *Adv. Chem. Phys.*, 1988, **70**, 489.
49. Grote, R. F. and Hynes, J. T., *J. Chem. Phys.*, 1980, **73**, 2715.
50. Bagchi, B. and Oxtoby, D. W., *J. Chem. Phys.*, 1983, **78**, 2735.
51. Bagchi, B. and Fleming, G. R., *J. Phys. Chem.*, 1990, **90**, 9.
52. Sundstrom, V. and Gilbro, T., *J. Chem. Phys.*, 1984, **81**, 3463.
53. Todd, D. C. and Fleming, G. R., *J. Chem. Phys.*, 1993, **98**, 269.
54. Nokowa, L., Schwarzer, D., Troe, J. and Schreder, J., *J. Chem. Phys.*, 1992, **97**, 4827.
55. Michel-Beyerle, M. E., (ed.), *Reaction Centres of Photosynthetic Bacteria*, Springer, Berlin, 1991.
56. Chandler, D., Gehlen, J. N. and Marchi, M., in *Ultra-fast Reaction Dynamics and Solvent Effects* (eds Gauduel, Y. and Rossky, P. J.), AIP Conf. Proc., Vol. 298, 1993.
57. Dobler, J., Zinth, W., Kaiser, W. and Desterhelt, D., *Chem. Phys. Lett.*, 1988, **144**, 215.
58. Honig, B., Ebery, T., Callender, R. H., Dinur, U. and Ottolenghi, M., *Proc. Natl. Acad. Sci. USA*, 1979, **76**, 2503.
59. Austin, R. H., Beeson, K. W., Eisenstein, L., Frauenfelder, H. and Gunsalus, C., *Biochemistry*, 1975, **14**, 5355.
60. Frauenfelder, H. and Wolynes, P. G., *Science*, 1985, **229**, 337.

ACKNOWLEDGEMENTS. The work reported here was carried out in collaboration with my students. I would particularly like to thank Srabani Roy, N. Gayathri, Dr A. Chandra and Dr N. Nandi for help and discussions. I thank Prof. P. Balaram and Ms N. Gayathri for a critical reading of the manuscript. It is a pleasure to thank Prof. Paul Barbara, Prof. Graham Fleming, Prof. Mark Maroncelli, Prof. A. Nitzan, Prof. Iwao Ohmine, Prof. Masonari Tachiya and Prof. Keitaro Yoshihara for many correspondences and discussions. I also thank Dr K. Bhattacharya, Dr P. K. Das, Dr S. Umapathy, Prof. D. Mukherjee and Prof. V. Krishnan for helpful discussions. Prof. C. N. R. Rao has always been a source of encouragement during the course of the work. This work was supported, in part, by the Council of Scientific and Industrial Research, India.

Received 26 April 1995; accepted 2 May 1995

Membrane current and potential change during neurotransmission in smooth muscle

Rohit Manchanda

School of Biomedical Engineering, Indian Institute of Technology, Powai, Bombay 400 076, India

Smooth muscle cells are electrically coupled to one another in a syncytium, and this renders their electrophysiology during neurotransmission strikingly different from that at other synapses. The postjunctional depolarizing responses of sympathetically innervated smooth muscle such as the vas deferens, particularly, the excitatory junction potentials (EJPs), possess intriguing properties which for several years have resisted explanation. A principal issue has been the temporal relationship of transmitter-generated membrane current to the resulting potential change, which seems to differ depending upon whether transmitter release is spontaneous or is nerve-stimulation-evoked. Accordingly, smooth muscle electrical properties appear to change with different patterns of transmitter release. Until some years ago this relationship was an area of uncertainty, firstly because transmitter-activated membrane current could not be measured directly and secondly because intracellular membrane potential measurements gave rise to conflicting results. Many of the uncertainties have now been resolved with refinements in techniques of measurement that have allowed membrane current time course during neurotransmission to be estimated. As a result, our understanding of smooth muscle electrical properties has been clarified and deepened. These developments are outlined in this review, and it is shown how our comprehension of neurotransmission has at every stage been influenced strongly by the techniques adopted for investigation.

THE electrophysiology of neurotransmission from autonomic nerves to smooth muscle has consistently been more challenging to study than somatic neuromuscular transmission. To a large extent this owes to the technical difficulties involved in making electrical recordings from smooth muscles. It is relatively problematic to record intracellularly from individual smooth muscle cells, which are just 2–5 μm in diameter, in contrast to skeletal muscle cells, whose diameters can be 50–150 μm. One requires to use microelectrodes with very small tip diameter (≤ 0.2 μm) and high tip impedance (50–200 MΩ). Further, when operating at these submicroscopic ranges the recording apparatus is exquisitely sensitive to mechanical vibration. Finally, when recordings are obtained their interpretation in terms of tissue electrical properties often rests on uncertain ground. This is because smooth muscle, as outlined below, is a complex electrical syncytium with cells interconnected to one another in three dimensions[1,2], and their electrical properties are poorly understood. Hence, even the primary depolarizing electrical responses of smooth muscles during neurotransmission, the excitatory junction potentials (EJPs), are not satisfactorily understood. In many cases these responses continue to be explored at the phenomenological rather than at the analytical level. By contrast, the processes involved in transmission at skeletal muscle have been rather thoroughly elaborated[3,4], starting from the end-plate potential down to the gating properties of the transmitter-activated receptor–ion-channel complex and the explanation of the end-plate currents in terms of channel function[5,6]. However, in recent years considerable progress has been made in our understanding of the generation of smooth muscle EJPs as well as the membrane currents underlying them. Some of this owes to the application of novel electrical recording methods to smooth muscle[7,8], and some to theoretical work. Several interesting observations and hypotheses have arisen, and it will be the aim of this article to review these developments. To set them in context, the historical development of ideas on the electrical properties of smooth muscles during neurotransmission will first be outlined. This will include a discussion of the 'cable' properties of smooth muscle, and the factors determining current spread and membrane potential development during EJPs. Since the field is now quite extensive, attention will centre on the properties of sympathetically innervated smooth muscle, particularly the vas deferens, which today provides an interesting and fruitful territory of research. Action potentials in smooth muscles will not be discussed, arising as they do subsequent to junctional transmission. For a survey of this topic the reader is referred to the detailed reviews of Tomita[2] and Huizinga[9].

Syncytial nature of smooth muscle and its innervation

Smooth muscle cells are well known to be electrically interconnected to one another to form what is known as a 'three-dimensional syncytium'[10–14]. Bozler[15] in 1948 was the first to suggest the syncytial nature of smooth muscle. His hypothesis was validated by the demonstration that injected current can flow, and membrane potential changes can be recorded, along smooth muscle tissues over distances which are considerably greater than the average individual cell length. The spread of current and potential occurs presumably through intercellular pathways of low resistance[2,10,16,17]. The syncytial organization of smooth muscle cells renders their electrophysiology in many respects different to that of other kinds of cells, e.g. skeletal muscle fibres, which are electrically isolated from their neighbours.

a. Autonomic–smooth muscle *b*. Somatic–skeletal muscle

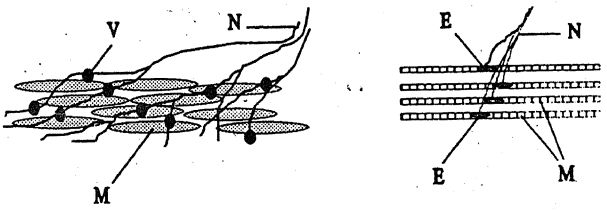


Figure 1. Comparison of patterns of innervation at the autonomic (*a*) and skeletal (*b*) neuromuscular junctions. N: nerve fibre; M: muscle cell; E: end-plate; V: varicosity. Smooth muscle cells are shown relatively shortened in length (e.g. in relation to varicosity diameter) for illustration.

The pattern of innervation of smooth muscle by autonomic nerves also differs from that of skeletal muscle by somatic nerves. As shown in Figure 1 *a*, the innervation of many smooth muscle organs is of the 'distributed' kind, that is, each cell may receive input from varicosities on more than one terminal nerve fibre and any terminal fibre may innervate several cells[14]. The result is a mesh of innervation known as the 'autonomic ground plexus', varicose in appearance, in which the points of transmission are thought to be the periodic axonal swellings, the varicosities. Some cells are innervated closely by the varicosities (contact distance < 20 μm) and others not[14,18]. In contrast to this meshwork, each skeletal muscle fibre receives a well-defined, localized innervation from a somatic motoneurone that forms a discrete 'end-plate' (Figure 1 *b*) where neurotransmission occurs[3–5]. Each skeletal muscle fibre is supplied by the terminal branch of only one somatic motor axon, and usually has just one end-plate.

As a consequence of the electrical interconnections between smooth muscle cells and of their complex innervation, any membrane potential change (such as an EJP) recorded in a particular cell will reflect not simply the response of that cell alone, but also the response, passively propagated to it, of neighbouring cells. Spatial and temporal summation of its own response may occur with the responses of neighbouring cells[11]. The extent of this 'pick-up' will depend largely on the electrical prop-

erties of the interconnecting pathways between the cells. For these reasons it is problematic to analyse electrical events at the level of individual smooth muscle cells in the kind of detail that has characterized the investigation of skeletal muscle, in which the electrical responses of an individual muscle fibre are independent of the responses of neighbouring cells.

Passive electrical properties

The passive electrical properties of a cell or a tissue, such as its membrane resistance and capacitance, and its length and time constants, are important determinants of the electrical response of the cell to any given input, since they determine its impulse response[19–21]. Thus, the amplitude, time course and spatial spread of a junction potential – be it the EPP in skeletal muscle, the excitatory or inhibitory postsynaptic potentials in neurones, or an EJP in smooth muscle, all of which are produced by neurotransmitter-activated transmembrane current – will depend largely on the passive properties. So will action potential properties such as conduction velocity[22,23].

Junction potentials usually have relatively rapid rising phases and slower, usually exponentially decaying, falling phases (Figure 2 *a*). In many kinds of cells the falling phase of the junction potential decays with a time constant that is equal to the cell's membrane time constant, τ_m. This indicates that the junction potential decay is a purely passive electrical process. The active part of the process, i.e. the neurotransmitter-activated membrane current, is relatively brief. Its duration is often of the order of the rise time of the junction potential (Figure 2 *b*). Transmitter-activated membrane current serves to discharge the membrane capacity and produce the rising phase of the junction potential, with the subsequent decay of the potential following the time course of passive recharging of the membrane capacity, at a rate determined by τ_m (refs. 5, 24). Thus, the time constant of decay, τ_{decay}, of the junction potential equals τ_m (Figure 2).

These properties are in many respects similar to those of potential changes across the capacitor in a parallel resistance–capacitance circuit in response to the flow of current through it. Based on this general principle, a standard method of evaluating the factors governing the time course of a junction potential has been to estimate τ_m and compare it with τ_{decay} of the junction potential. If the two values are similar, then the membrane currents and potential are held to behave as outlined above. If τ_{decay} exceeds τ_m, other factors – such as a duration of transmitter action greater than τ_m – have to be considered[25].

An approach such as this should in principle have revealed readily the elements of junctional processes in smooth muscle. As ever, though, smooth muscle presents bedevilling complications. One is that in a

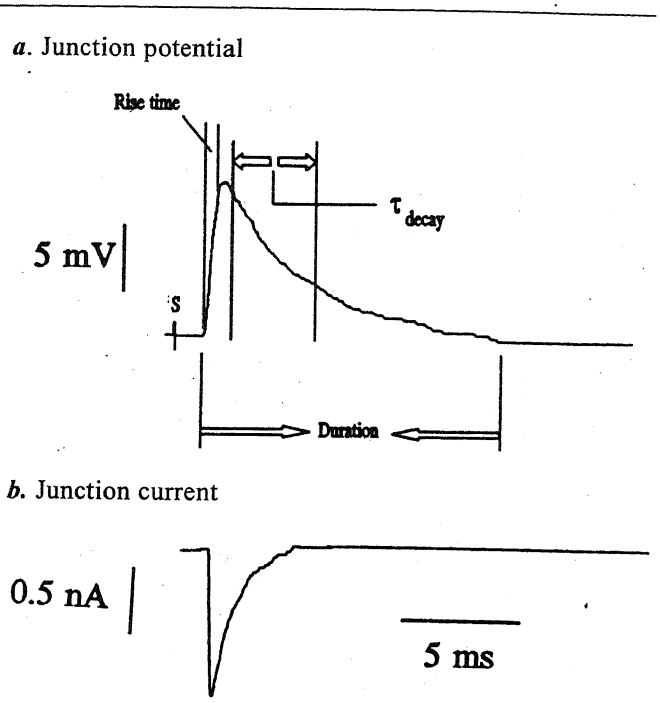


Figure 2. A typical junction potential (*a*) and its underlying junction current (*b*) as observed at a variety of synapses. S: stimulation artefact. Parameters of analytical interest are indicated. Note that membrane potential change is prolonged compared to membrane current, because of the capacitive properties of the cell membrane. The downward current trace indicates inward membrane current. Typical amplitudes and time courses observed at the skeletal neuromuscular junction are indicated by the calibration bars; however (as will be seen in subsequent figures), values may vary depending upon the cell type under investigation. It should also be noted that if junction current time course were measured using extracellular recording, its amplitude calibration would be in volts rather than in amperes.

syncytium, τ_m itself varies under different conditions of determination. The estimation of τ_m in simpler structures (those approximating linear one-dimensional cables, such as skeletal muscle fibres and axons) has been relatively straightforward. One injects rectangular pulses of constant current at a point inside the cable and measures the resulting membrane potential changes (the subthreshold or electrotonic responses) at increasing distances from the point of injection[3,5]. τ_m, and other passive parameters such as membrane resistance, capacitance and length constant, λ_m, can then be evaluated using the well-known properties of the cable equation[21].

By contrast, current injection at a point in the smooth muscle syncytium (i.e. intracellularly in a single cell) results in no detectable membrane potential change even in its immediate vicinity[11,26]. The reason for this is thought to be a very rapid spatial dissipation of injected current through the multiple shunts offered by intercellular electrical pathways in three dimensions, so that no appreciable fraction of it flows to any neighbouring area. Thus, smooth muscle does not exhibit cable properties with intracellular current injection. However, when it is subjected to current injection through large

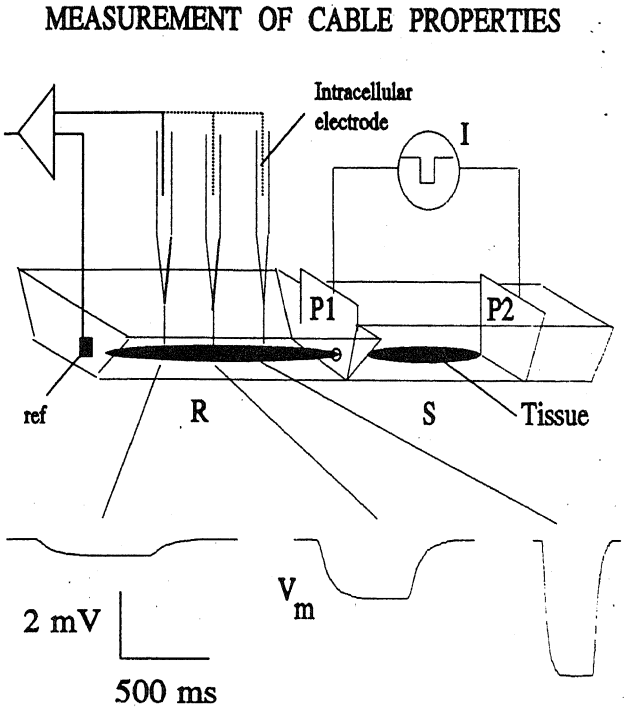


Figure 3. Partition stimulation chamber used for investigation of smooth muscle cable properties. R: recording chamber; S: stimulating chamber; P1, P2: metal plates for stimulation and partition; I: current source; V_m: membrane potential change at different distances from plate P1, recorded intracellularly (with reference to indifferent electrode 'ref'). The scale is exaggerated: V_m responses may be recorded up to only a few mm from P1. Also there would be a tight fit between the tissue and the hole in P1. Amplitude and time scales are appropriate for guinea pig vas deferens. Adapted from ref. 16.

extracellular electrodes, smooth muscle does, surprisingly, start to exhibit cable properties, so that membrane potential changes are recorded at a distance from the area of current injection. Particular mention may be made of the 'partition stimulation' method first employed by Tomita[10,11] and Abe and Tomita[16]. The experimental arrangement is shown in Figure 3. The cylindrical smooth muscle organ is drawn through a circular hole of matching diameter in a metallic plate (P1) which forms a plane at right angles to the longitudinal axis of the tissue. The plate also separates the organ bath in which the tissue is placed into two compartments electrically insulated from each other. Current is passed extracellularly between this plate and another plate electrode (P2) in the stimulating compartment (S), while in the recording compartment (R) membrane potential responses are measured intracellularly. As shown in Figure 3, passive (electrotonic) responses can now be detected in tissue cells up to a few millimetres from the plate, and a quantitative evaluation of these responses indicates that the organ now behaves as a one-dimensional cable. An explanation for this change of properties is that when external plate electrodes are used to polarize the smooth muscle cells, the tissue is rendered more or less isopotential in its two radial axes at the stimulation plate. Hence, injected current can only spread in one direction – along the longitudinal aspect of the tissue – and the tissue acquires cable-like properties in this direction[10,11,17]. Some of these properties are: (i) the peak values of the membrane potential responses to external stimulation decay exponentially with distance from the polarizing plate; (ii) the temporal properties of the membrane potential changes at different distances can be described by linear one-dimensional cable equations.

The electrotonic responses thus measured can be used to estimate passive electrical properties such as length constant and time constant of the tissue[11,27]. In the case illustrated (hyperpolarizing current injection) the change of membrane potential (V_m) with respect to both time (t) and distance (x) is given by[16,19]

$$V_m(X,T) = V_m(X=0, T=\infty)\frac{1}{2}\left\{e^{-X}\left[1+\mathrm{erf}\left(\frac{X}{2\sqrt{T}}-\sqrt{T}\right)\right] + e^{X}\left[1-\mathrm{erf}\left(\frac{X}{2\sqrt{T}}+\sqrt{T}\right)\right]\right\}, \quad (1)$$

where $X = x/\lambda_m$ (here $0 < X < \infty$) and $T = t/\tau_m$, x being the distance from the stimulating electrode and t the time from start of stimulation; erf is the error function, defined as

$$\mathrm{erf}(y) = \frac{1}{2\sqrt{\pi}}\int_0^y e^{-\omega^2}\,d\omega. \quad (2)$$

By comparing the prediction of equation (1) with the observed responses, one can verify the correctness of the values of the length constant and time constant obtained experimentally, and thus estimate the passive membrane properties.

Based on the responses to intracellular and extracellular polarization, various suggestions, often conflicting, have been made about the factors determining the time courses of EJPs, and the relationship of membrane potential change to membrane current, in smooth muscle. These hypotheses are reviewed below.

Smooth muscle junction potentials and their analysis

Two kinds of excitatory potentials are recorded at the synapse between the terminals of the vas deferens nerve postganglionic to the hypogastric nerve and the smooth muscle cells of the vas[28,29]: the evoked and spontaneously occurring junction potentials (the EJP and SEJP, respectively, Figure 4 *a*). EJPs and SEJPs have been recorded in several autonomically innervated smooth muscles, but the most detailed work has been done on junction potentials at the sympathetic neuroeffector junction, particularly in the rodent vas deferens, and in arteries and arterioles.

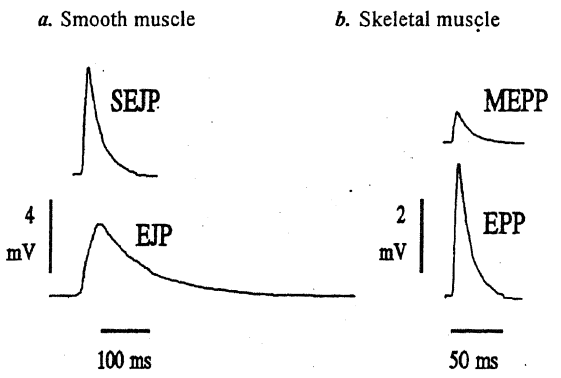


Figure 4. Spontaneous and stimulation-evoked junction potentials seen in smooth (*a*) and skeletal (*b*) muscle. Spontaneous potentials: SEJP, MEPP. Evoked: EJP, EPP. In smooth muscle the SEJP is briefer than the EJP but can be as large as the EJP. In skeletal muscle the MEPP follows the same time course as the EPP but has a smaller amplitude. Note the different time scales in *a* and *b*.

The SEJP occurs in the absence of nerve stimulation and is thought to be caused by the spontaneous release of a quantum of transmitter from a neuronal varicosity. The EJP, on the other hand, is a stimulation-evoked event and is thought to be caused by the release of transmitter throughout the neuronal ground plexus following nerve stimulation[30,31]. Qualitatively, these responses are analogous to the miniature and the stimulation-evoked end-plate potentials (MEPPs and EPPs) observed at the skeletal neuromuscular junction, produced, respectively, by the release of one quantum of transmitter (spontaneously) or of several quanta (following nerve stimulation)[32] (Figure 4 *b*). However, there are important differences, as illustrated in Figure 4. In relation to time courses the most striking one is that whereas the MEPP and the EPP have similar time courses (Figure 4 *b*), the SEJP and the EJP differ by a factor of 5–10 in theirs, the SEJP being briefer[28,33] (Figure 4 *a*). SEJPs usually have durations of 100–150 ms and τ_{decay} of 20–50 ms, whereas EJPs have durations of 0.6–1 s and τ_{decay} ranges from 200 to 400 ms.

Another difference lies in relative amplitudes. Skeletal muscle MEPPs have amplitudes that are normally distributed about the range 0.5–1 mV, and the EPP amplitude is an integral multiple of this quantal level; EPPs are thus always as large as, and normally much larger than, MEPPs[32]. By contrast, in smooth muscle both EJP and SEJP amplitudes are continuously variable from the lowest to the highest levels (0–30 mV), EJP amplitude being graded with stimulation intensity and SEJP amplitudes varying randomly. There is no obvious quantal relationship between the two events, and SEJPs can often be larger than EJPs[28,33] (Figure 4).

Since MEPPs and EPPs have similar time courses, it would appear that the passive electrical properties of skeletal muscle cells remain invariant regardless of whether transmitter release is spontaneous or stimulation-evoked. However, in smooth muscle cells, since the SEJP and EJP follow very different time courses, the passive properties appear to change depending upon the pattern of transmitter release, i.e. whether it is spontaneous and localized to a varicosity or is evoked by nerve stimulation and, therefore, involves several varicosities throughout the ground plexus[11,13].

Using the principle of comparison of τ_m with τ_{decay} of the junction potentials, the decays of the MEPP as well as of the EPP of skeletal muscle were shown to be determined by the passive properties of the skeletal muscle cell membrane. However, in initial studies the decays of the SEJP and EJP in the vas deferens were suggested by various workers to reflect not τ_m but the time course of membrane current, indicating a prolonged duration of action of transmitter at the autonomic neuroeffector junction.

Experimental studies

The SEJP. For the SEJP, evidence for prolonged transmitter action came mostly from studies on membrane potential responses to intracellular stimulation. If the SEJP results from localized transmitter action at a point in a three-dimensional syncytium, then the electrical properties of the smooth muscle membrane during intracellular current injection may be assumed to reflect its properties during an SEJP. A number of workers[11,34–36] have studied the membrane potential responses of smooth muscle cells to intracellular current injection with current-passing microelectrodes. In the guinea pig and mouse vas deferens this response was usually observed to have a time constant of 2–7 ms, which is up to an order of magnitude higher than the time constant of decay of the SEJP (range 20–35 ms). It could even be more rapid than the rising phase of the SEJP, which occupies between 5 and 15 ms.

Thus, the time constant of the membrane in response to local application of current in the syncytium was much too brief to account for the time constant of decay of the SEJP. This excluded the possibility that the passive membrane properties of the vas deferens determined the decay of the SEJP. Hence, for the SEJP it was concluded that transmitter action might continue throughout its duration[10,11].

The EJP. Suggestions about the factors governing EJP time course came primarily from estimation of smooth muscle passive properties using the partition–stimulation method. From the passive responses described by equation (1), Tomita[11] estimated the length constant (λ_m) of the tissue to be about 1.5–2.5 mm. The membrane time constant τ_m of the smooth muscle cells was estimated to be about 100 ms.

The time constant of the declining phase of the EJP (200–400 ms) is significantly greater than this value of τ_m. Hence, it appeared that τ_m was too brief to account for the falling phase of the EJP, rather like the relationship of τ_m during intracellular polarization in relation to the SEJP. It was concluded[11,17,26] that the EJP too might reflect the time course of ongoing transmitter action rather than that of the passive recharging of the membrane capacity.

Thus, until the mid-1970s there seemed to be a consensus that transmitter-activated membrane current during both the SEJP and the EJP might persist throughout the duration of each category of junction potential. This contrasted with the brief time course of membrane current underlying the MEPP and the EPP at the skeletal end-plate. In the last 10–15 years, however, much evidence has been presented which suggests a somewhat different picture for the EJP. Let us consider these developments.

The most pertinent experimental analysis is that of Bywater and Taylor[27], who reinvestigated the passive electrical properties of the guinea pig vas deferens. They showed that the estimation of membrane time constant depended critically on the respective lengths of the tissues placed in the recording and stimulating compartments of the partitioned chamber. In particular, the calculated time constant was shown to be a considerable underestimate of its true value if less than three length constants of the tissue were placed in the stimulating compartment.

Obviating these sources of error, τ_m was estimated[27] with external polarization to be about 270 ms, a value significantly different from the previous estimate (100 ms)[11]. The length constant λ_m (~860 μm) was also shown to be smaller than the earlier estimate[11] of 1500–2600 μm.

Unlike the earlier value, the new estimate of τ_m coincided with the time constant of decay of the EJPs in the guinea pig vas deferens. The decay of the EJP, therefore, appeared to be dictated by the passive membrane properties of the smooth muscle cells. This would indicate a brief duration of transmitter-activated current underlying the EJP. Blakeley and Cunnane[30] and Bywater and Taylor[27] suggested that transmitter action underlying both spontaneous and evoked junction potentials corresponded in duration to the SEJP.

Theoretical studies

Smooth muscle syncytial properties have also been investigated from a theoretical standpoint[10,11,13]. It was shown that if current were injected from a point source into the syncytium (as for the SEJP) then, because of intercellular electrical coupling, the current would dissipate away from the point of injection very rapidly, both spatially as well as temporally[11]. As a result, the time course of the membrane potential change at or near the point of current injection would be similar to that of the current itself. On the other hand, when current of the same time course is injected uniformly throughout the syncytium (as for the EJP), spread of current is restricted and the potential change that develops then is prolonged, lagging considerably behind the current[13].

Purves[13] obtained the time course of the EJP by convolving the impulse response, $h(T)$, of an isopotential circuit with the input, the transmitter-activated current $I(T)$:

$$h(T) = \frac{Q_0}{c_m} e^{-T}, \tag{3}$$

$$I(T) = \alpha^2 T \exp(-\alpha T), \tag{4}$$

where Q_0 is the charge deposited instantaneously on, and c_m the membrane capacity of, unit volume of tissue, T is the normalized time (t/τ_m) and α is a driving function which, as used in equation (4), generates waveforms very similar in shape to synaptic currents observed at a variety of synapses[13,23]. Then the membrane potential response $V(T)$ is given by $V(T) = h(T)*I(T)$, the asterisk indicating the convolution operation. On solving the convolution integral we get the essential part of $V(T)$ as

$$V(T) = \alpha^2 e^{-\alpha T}\left\{\frac{e^{T(\alpha-1)}-1}{(\alpha-1)^2} - \frac{T}{\alpha-1}\right\}. \tag{5}$$

On evaluation, $V(T)$, corresponding here to the EJP, turns out to be considerably prolonged compared to the injected current $I(T)$.

Purves[13] also derived an expression for the voltage change during the SEJP. In this case the impulse response $h(T)$ is given by

$$h(T) = \frac{Q_0 r_i}{8(\pi T)^{3/2}\lambda\tau} e^{[-(R^2/4T)-T]}, \tag{6}$$

r_i being the intracellular resistance per unit volume and R the normalized radial distance.

Equation (6) can be convolved with the expression for transmitter-produced membrane current (equation (3)) and the voltage response derived. It was shown that the resulting voltage change is much briefer than that predicted for the isopotential EJP and, in fact, follows closely the time course of the current itself[13]. Hence, this theoretical work indicated that whereas the EJP is prolonged compared with its underlying membrane current, the SEJP follows the time course of the current.

Recent investigations

The foregoing conclusions can be unambiguously tested only by direct observation of the junction currents that underlie the potentials to establish the relationships between their time courses. One way to record junction

current is to perform a voltage clamp experiment, which has been done on several types of electrically isolated cells, including skeletal muscle fibres and neurones[4,24,37]. However, it is difficult normally to obtain an adequate voltage clamp of syncytial, three-dimensional tissues such as smooth muscle[38]. This is because, as previously described, smooth muscle cells have unsuitable geometries and are electrically interconnected.

Extracellular recording. A second convenient method of estimating membrane current time course is to record extracellularly the potential changes caused by the current sinks created focally at points of transmitter action[39]. Until the late 1980s this method could not be successfully applied to the autonomic-nerve–smooth-muscle junction because of technical problems[7,40,41]. However, in the past few years attempts at extracellular recording have been successful and, therefore, it has been possible to estimate the time course of membrane current underlying the EJPs[7,41–43]. The extracellular electrode used in the initial studies had a large tip diameter (20–50 μm) compared to the intracellular electrode, and was applied to the surface of the vas deferens with slight suction. This allowed the recording of negative-going extracellular potential changes of 20 μV and greater in amplitude. Since these events reflect the time course of membrane current, they were named excitatory junction currents (EJCs)[7].

Simultaneous intracellular and extracellular recordings of the time courses of junction potentials and currents have been obtained[8,33] and have given rise to some conclusive and interesting observations. In the experiments in which the spontaneous events were recorded, cells within a short distance (up to 100 μm) of the rim of the extracellular electrode were probed with the intracellular microelectrode[8,44]. This was necessary because, as mentioned above, the spatial decay of spontaneous potentials is very rapid and the same event could be detected with two electrodes only if they were very close to each other. The experiments unequivocally showed that when the same electrical event was recorded at both the electrodes[44], the time courses of the SEJP and the SEJC (i.e. the potential change and the current flow) were identical[8]. This is shown in Figure 5 for three pairs of SEJPs and their underlying SEJCs, and it is noteworthy that even for very different shapes of the events in Figure 5 *a*, *b* and *c*, potential change and current follow the same time course. The result strongly suggests that the membrane capacity is not charged significantly by the current induced by the action of a single quantum of transmitter, since if it were, membrane potential change would have lagged behind the current.

The nerve-stimulation-evoked EJP and the time course of its underlying EJC were also recorded simultaneously[8,33]. In this case it was observed that the membrane currents underlying the EJPs were brief, having

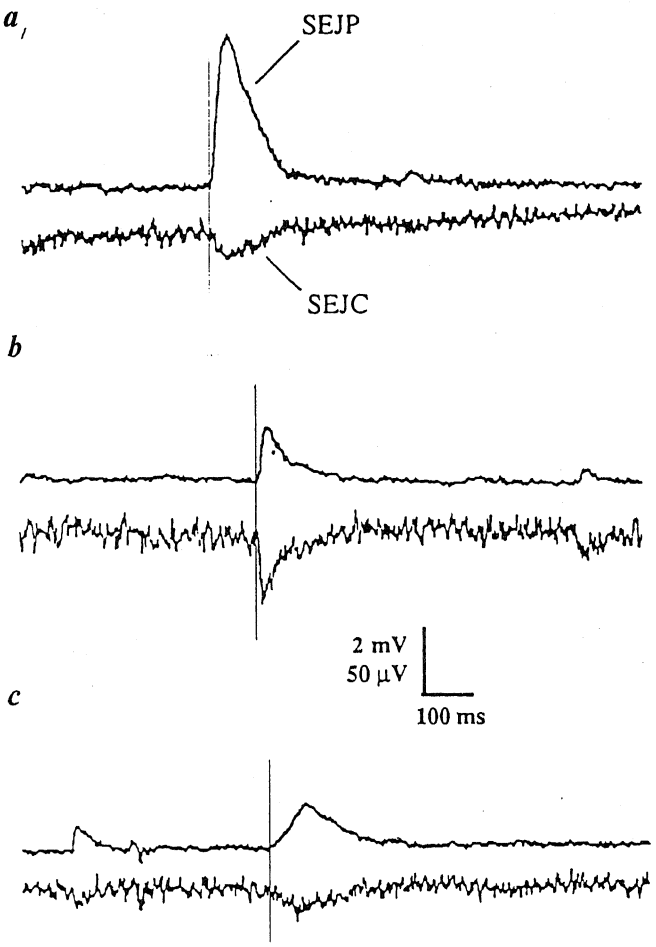


Figure 5. Spontaneous electrical events in smooth muscle (mouse vas deferens) recorded simultaneously with intracellular and extracellular electrodes (upper and lower traces, respectively, in *a*, *b* and *c*). Note the close correspondence between the time courses of membrane potential change (SEJP) and junctional current (SEJC) even for different shapes of the events: the almost linear decays in *a*, the biexponential decays in *b*, and the prolonged rise times in *c*. The reason for the somewhat distorted shapes is probably the mechanical pressure exerted by the extracellular electrode on the tissue[44]. Voltage scale: 2 mV for intracellular, 50 μV for extracellular traces.

durations (100–150 ms) and time constants of decay (25–40 ms) no greater than of the SEJCs (Figure 6). Hence, the prolonged time constant of decay of the EJPs (250–350 ms) cannot be attributed to prolonged transmitter action, owing, for instance, to slow inactivation or to physical barriers to diffusion[1]. It must instead be attributed to the passive decay of charge from the smooth muscle cells following spatially distributed injection of current into the tissue that changes it from a three-dimensional syncytium into one that is presumably at isopotential[13,33]. If this indeed is the case, the theoretical isopotential voltage response of the tissue (given by equation (5)) should predict successfully the experimental data. Evidence for this is presented in Figure 7. EJPs and EJCs were recorded simultaneously and 80 of each were averaged[33,44] (Figure 7 *a*). The EJC was

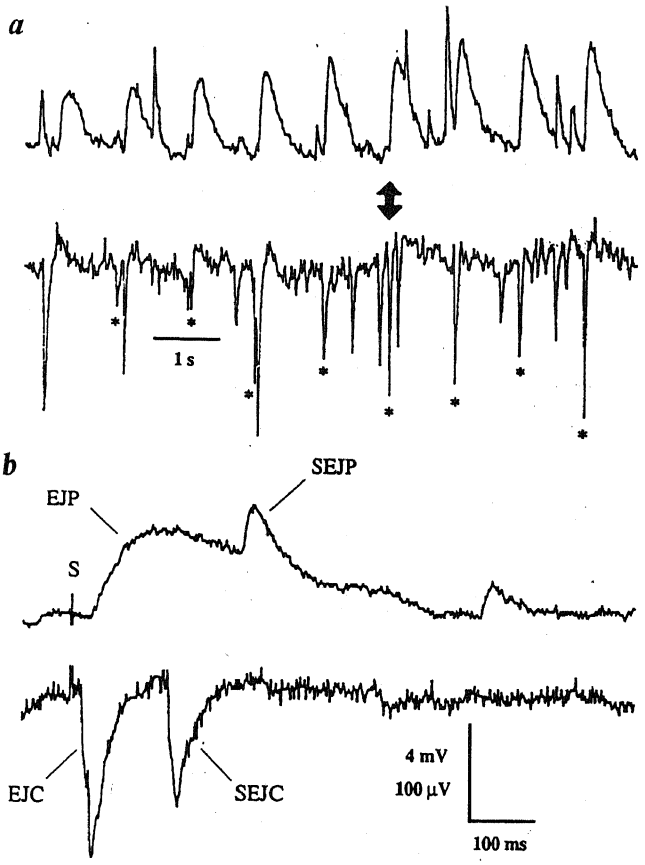


Figure 6. Stimulation-evoked electrical events in smooth muscle (guinea pig vas deferens) recorded simultaneously with intracellular and extracellular electrodes. A series of 9 EJPs is shown in *a*, the underlying EJCs marked with asterisks. In *b*, one of the paired events in *a* (marked with an arrow) is shown on an expanded time scale. Note the brief time course of the EJC compared to the EJP. Other deflections in *a* are spontaneously occurring SEJPs and SEJCs; one of each is indicated in *b*. Voltage scale: 4 mV for intracellular, 100 μV for extracellular traces. S: stimulation artefact.

shown to match the form of the theoretical junction current given by equation (3) (Figure 7 *b*). When this current was used as the input to an isopotential system and the voltage response computed, the response agreed very closely with the experimental EJP (Figure 7 *c*). In these recordings, therefore, the tissue appears to be at isopotential during the EJP[13,33]. In further support of this conclusion is the observation that under conditions of field stimulation of the tissue, the estimated membrane time constant of these cells (270–300 ms)[27] is very similar to the time constant of decay of the EJPs.

Bennett *et al.*[45] have carried out a theoretical analysis of the signals that one would expect to obtain with an extracellular surface electrode of the kind used in the studies mentioned above. Their predictions are in very good accord with the EJCs experimentally recorded earlier[8,41,43], and place confidence in the validity of the new measurements.

The quantal relation between the EJP and SEJP

Although the time courses of the SEJP and the EJP are widely different, the events have similar ranges of amplitudes, from just-discernible to threshold levels for action potential generation. The range of amplitudes of SEJPs can be explained by the effect of electrotonic decrement on events generated in a syncytium at varying radial distances from the recording electrode[11,13]. Events generated very close to the recording electrode are recorded as large SEJPs, and those at larger distances as smaller SEJPs. However, the EJP seems not to be composed of an integral multiple of SEJPs, as the quantal hypothesis in its simplest form (derived from the properties of the EPP and MEPP at the skeletal neuromuscular junction[32]) would suggest. This feature of sympathetic neurotransmission, by itself a separate important area of investigation, merits a brief comment. The problem here is that the EJP and the SEJP arise under different prevailing electrical conditions of the tissue (as detailed above). The EJP for most of its time course is a passive electrical event, and only its rising phase is actively determined by the underlying junction current (see Figures 6 and 7). So, a resolution of transmitter release events following nerve stimulation, and electrical identification of sites of release, would require investigation of its rising phase, in contrast to the decay phase, which throws light on the tissue's electrical properties.

Detailed examination of the rising phases of EJPs has revealed that they are not entirely smooth but are marked by inflexions, and that both the slope and the time of occurrence (latency) of these inflexions differ from one EJP to the next[30]. The first time differentials (dV/dt) of the rising phases, therefore, contain sharp peaks corresponding to the inflexions, and are termed 'discrete events' (DEs). DEs were postulated to be indications of the activity of individual transmitter release sites close to the recording electrode, the remainder of the EJP reflecting the background level of depolarization in the surrounding tissue. Rising phases of SEJPs were similarly differentiated to give rise to spontaneous DEs (SDEs).

Using DEs it has been possible to 'fingerprint' different release sites around the intracellular recording electrode by virtue of the differences in amplitudes, time courses and (for evoked events) latencies of the DEs[31]. Evidence from these studies suggests that stimulation-evoked release from an individual electrophysiologically identified release site (possibly a varicosity) is highly intermittent, and monoquantal. That is, in a train of stimuli, only one in several stimuli activate a particular release site. And when a site is activated, the resultant evoked DE corresponds exactly to a particular SDE observed in that cell[31]. It is not an integral multiple of the SDE. Thus, evoked release must be monoquantal if the

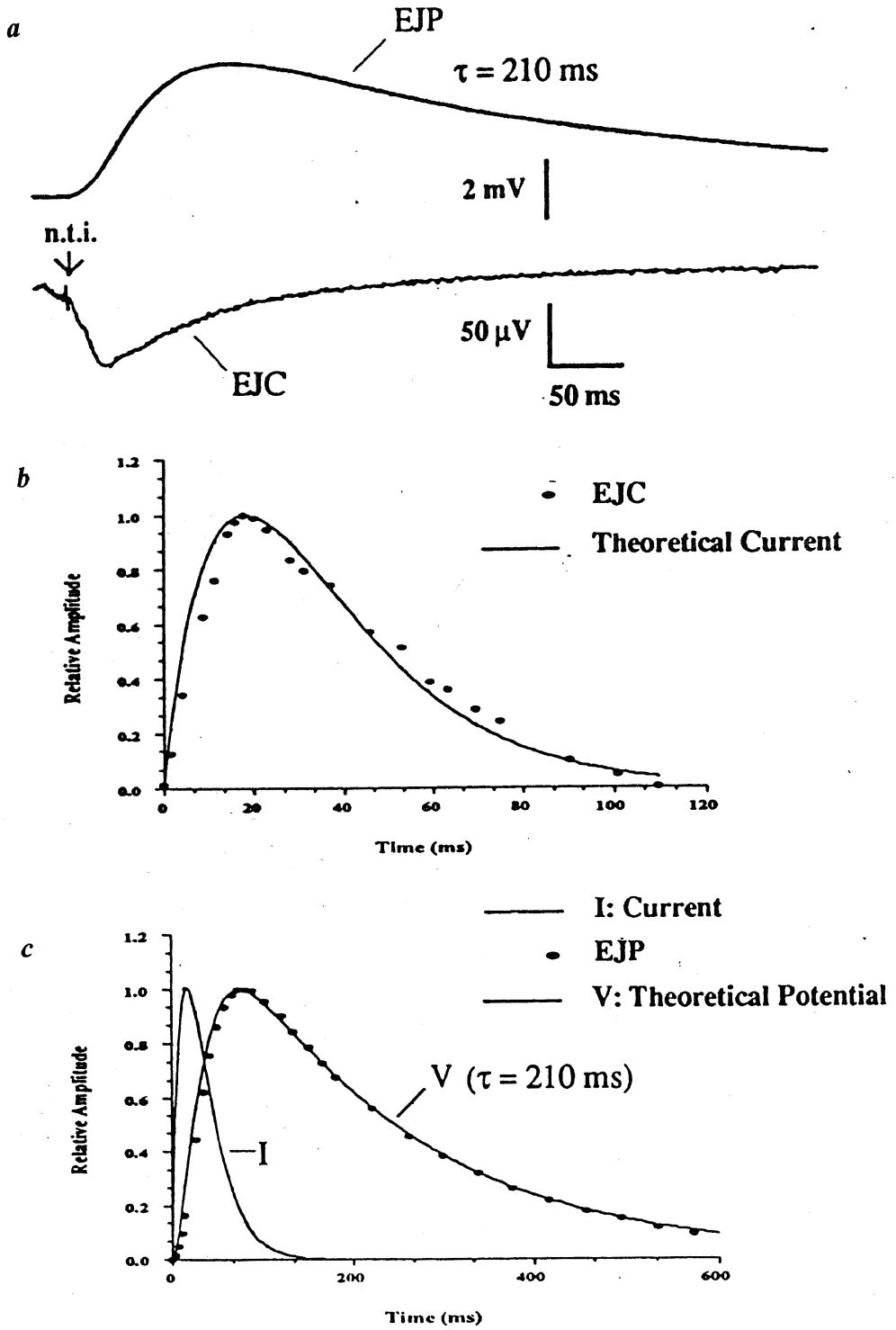


Figure 7. Reconstruction of the EJP from the EJC based on Purves' model[13, 44]: *a*, averages of 80 EJPs and EJCs simultaneously recorded (n.t.i. is the nerve terminal impulse recorded extracellularly); *b*, plot of the EJC in *a* (filled circles) compared to theoretical EJC (smooth curve) generated by equation (4) of the text; *c*, plot of experimental EJP in *a* (filled circles) along with theoretical EJP (smooth curve) generated by Purves' model (equation (5) of text) using the underlying EJC (*I*) as the input. Note the agreement between experimental and theoretical time courses. In *b* and *c*, all events have been normalized with respect to their maximum amplitudes in order to facilitate comparison of time courses.

SEJP is held to be a monoquantal event (on which there is general consensus). Analysis of EJCs and SEJCs using the extracellular recording technique has confirmed this hypothesis of intermittent, monoquantal release from release sites[41–43,46]. However, it should be noted that during an EJP several release sites throughout the

ground plexus are activated, and in this sense the EJP is a multiquantal event. But because the activated release sites are spatially dispersed[46], in the EJP's rising phase only the contribution of one of the closest sites is usually recorded intracellularly as a DE, the contribution of more distant sites being attenuated by electrotronic factors. Since release from each of the sites is monoquantal, and the effect of no more than one of them may be recorded during a given EJP, transmitter release will not appear to be multiquantal. The EJP is graded with stimulation strength because at low strengths fewer intramural axons, and hence varicosities, are activated than at higher strengths, leading to different average levels of depolarization in the tissue. Hence, the EJP can be as small as the smallest SEJP, or as large as the largest, depending upon the stimulation intensity. For a more detailed account of this somewhat unusual quantal relation, the interested reader is referred to the relevant papers and reviews[30,31,41–43,46].

Summary

To summarize, theoretical predictions and experimental results are now in good agreement about the factors governing junction potential time courses in smooth muscle. Extracellular recordings have made it possible to measure the kinetics of junction currents. Simultaneous intracellular and extracellular recordings have established that whereas in the case of the SEJP membrane potential change follows the time course of underlying current, in the case of the EJP it is much more prolonged than the current. Spontaneous and evoked junction currents are similar in time course. These results can be explained in terms of the syncytial nature of smooth muscle by postulating redistribution of charge in a three-dimensional resistive network in the syncytium during the SEJP and in an approximately isopotential resistive–capacitative network during the EJP.

Support to this scheme has been lent by studies on other kinds of smooth muscle, notably vascular smooth muscle, in one particularly interesting case by demonstrating a corollary to the hypothesis. Hirst and Neild[47,48] described a guinea pig submucosal arteriolar preparation which was just one smooth muscle cell layer thick. By using short segments of arteriolar branches which were less than 0.35 length constants long, they showed that the arteriole behaved as a one-dimensional cable even for intracellular current injection. Under these conditions, the time constant of decay of not only the EJP[47] but also the SEJP[48] was shown to be similar in value to τ_m, and the time courses of the EJP and the SEJP were identical. Hirst and Neild[47] also estimated that the membrane currents underlying both the EJP and the SEJP were relatively brief. Hence, when an otherwise three-dimensional syncytium is rendered – by fortuitous physiological arrangement – unidimensional and cable-like, its electrical behaviour becomes similar to that of a skeletal muscle cell. Finkel *et al.*[49] succeeded in voltage clamping this one-dimensional arteriolar preparation with a single-microelectrode voltage clamp system and corroborated these suggestions. It is interesting that the time course of the conductance change in this vascular smooth muscle resembles the time course of the SEJP in the vas deferens[47–50]. In other vascular smooth muscle, where three-dimensional syncytial properties are exhibited, tissue electrical properties resemble those of the vas deferens[51].

Conclusions

In this review I have outlined the often tortuous development of ideas on smooth muscle electrical properties during neurotransmission. It is noteworthy at every stage that the methods adopted for determination of electrical properties have strongly influenced our theoretical picture. Thus, although we now have a hypothesis that consistently explains many observations and may have heuristic value, it might subsequently require modification in the face of new observations obtained with more refined techniques. Additional questions that require attention are:

(i) In some smooth muscles evidence from different lines of experimentation is in conflict regarding whether the cells are electrically coupled. In the vas deferens of the rat and the mouse it has not been possible to demonstrate cable-like properties using the partition stimulation method[36], indicating restricted or absent intercellular current flow, at least in the longitudinal direction. Yet, in these tissues the relationship between the time courses of the junction potentials and the currents is as outlined above[44], indicating effective electrical coupling. Such discrepancies need to be resolved.

(ii) What are the properties, in terms of unitary conductance and mean open time, of the neurotransmitter-activated ion channels which underlie the development of the SEJP and the EJP? Patch clamp studies have so far failed to provide the answer, as only whole-cell recordings have been possible[12]. An alternative approach to the problem might be to carry out fluctuation analysis of transmitter-activated membrane noise[5,37], which would provide indirectly, but reliably, the information of interest.

(iii) Not all EJPs have simple time courses. In many cells their decaying phases can be complex, being described by more than one time constant of decay[52], and one of these phases may be briefer than τ_m. Theoretical studies need to be done on how such EJPs might arise. This may perhaps best be begun by considering simple equivalent circuits for the tissue, intermediate in complexity between the three-dimensional syncytial and the isopotential cases, and estimating the voltage responses in these[53].

(iv) The identity of the subcellular structures that allow intercellular electrical coupling in smooth muscle is controversial. In other syncytial tissues, the structure generally accepted to underlie coupling is the 'gap junction'[54]. However, in the case of smooth muscle the evidence for the existence of gap junctions between cells is equivocal, as is their function when present[54,55]. It will be interesting to see whether a morphological substrate is soon established for the electrical coupling that is so strikingly evident functionally and that plays such an important role in determining tissue electrical properties under different conditions.

1. Bennett, M. R., *Philos. Trans. R. Soc. Londo*n, 1973, **B265**, 25–34.
2. Tomita, T., *Prog. Biophys. Mol. Biol.*, 1975, **30**, 185–203.
3. Katz, B., *Nerve, Muscle, and Synapse*, Mc-Graw Hill, New York, 1966.
4. Prior, C., Dempster, J. and Marshall, I. G., *J. Pharmacol. Toxicol. Methods*, 1993, **30**, 1–17.
5. Nicholls, J. G., Martin, A. R. and Wallace, B., *From Neuron to Brain*, 3rd edn, Sinauer, Sunderland, 1992.
6. Unwin, N., *Nature*, 1995, **373**, 37–43.
7. Brock, J. A. C. and Cunnane, T. C., *Nature*, 1987, **326**, 605–607.
8. Cunnane, T. C. and Manchanda, R., *Neuroscience*, 1989, **30**, 563–575.
9. Huizinga, J. D., *Can. J. Physiol. Pharmacol.*, 1991, **69**, 1133–1142.
10. Tomita, T., *J. Theor. Biol.*, 1966, **12**, 216–227.
11. Tomita, T., *J. Physiol.*, 1967, **189**, 163–176.
12. Tomita, T., *Jpn. J. Physiol.*, 1992, **58**, 1P–6P.
13. Purves, R. D., *J. Theor. Biol.*, 1976, **60**, 147–162.
14. Burnstock, G., *Experientia*, 1985, **41**, 869–874.
15. Bozler, E., *Philos. Trans. R. Soc. London*, 1973, **B265**, 3–6.
16. Abe, Y. and Tomita, T., *J. Physiol.*, 1968, **196**, 87–100.
17. Tomita, T., in *Smooth Muscle* (eds. Bülbring, E., Brading, A. F., Jones, A. and Tomita, T.), Edward Arnold, London, 1970, pp. 197–243.
18. Lavidis, N. A. and Bennett, M. R., *J. Auto. Ner. Sys.*, 1993, **45**, 87–100.
19. Hodgkin, A. L. and Rushton, W. A. H., *Proc. R. Soc. London*, 1946, **B133**, 444–479.
20. Katz, B., *Proc. R. Soc. London*, 1948, **B135**, 506–534.
21. Jack, J. J. B., Noble, D. and Tsien, R. W., *Electric Current Flow in Excitable Cells*, Oxford University Press, Oxford, 1983.
22. Fatt, P. and Katz, B., *J. Physiol.*, 1951, **115**, 320–370.
23. Jack, J. J. B. and Redman, S. J., *J. Physiol.*, 1971, **215**, 321–352.
24. Gage, P. W., *Physiol. Rev.*, 1976, **56**, 177–247.
25. Faber, D. S. and Korn, H., *Science*, 1980, **208**, 612–614.
26. Kuriyama, H., *J. Physiol.*, 1964, **175**, 211–230.
27. Bywater, R. A. R. and Taylor, G. W., *J. Physiol.*, 1980, **300**, 303–316.
28. Burnstock, G. and Holman, M. E., *J. Physiol.*, 1961, **155**, 115–133.
29. Burnstock, G. and Holman, M. E., *J. Physiol.*, 1962, **160**, 446–460.
30. Blakeley, A. G. H. and Cunnane, T. C., *J. Physiol.*, 1979, **296**, 85–96.
31. Cunnane, T. C. and Stjärne, L., *Neuroscience*, 1984, **13**, 1–20.
32. Del Castillo, J. and Katz, B., *J. Physiol.*, 1954, **124**, 560–573.
33. Cunnane, T. C. and Manchanda, R., *Neuroscience*, 1990, **37**, 507–516.
34. Hashimoto, Y., Holman, M. E. and Tille, J., *J. Physiol.*, 1966, **186**, 27–41.
35. Bennett, M. R., *J. Gen. Physiol.*, 1967, **50**, 2459–2470.
36. Holman, M. E., Taylor, G. S. and Tomita, T., *J. Physiol.*, 1977, **266**, 751–764.
37. Anderson, C. R. and Stevens, C. F., *J. Physiol.*, 1973, **235**, 655–693.
38. Bolton, T. B., Tomita, T. and Vassort, G., in *Smooth Muscle: An Assessment of Current Knowledge* (eds Bulbring, E., Brading, A. F., Jones, A. and Tomita, T.), Edward Arnold, London, 1981, pp. 47–63.
39. Del Castillo, J. and Katz, B., *J. Physiol.*, 1956, **132**, 630–649.
40. Holman, M. E., in *Smooth Muscle* (eds Bülbring, B., Brading, A. F., Jones, A. and Tomita, T.), Edward Arnold, London, 1970, pp. 244–288.
41. Brock, J. A. C. and Cunnane, T. C., *J. Physiol.*, 1988, **399**, 607–632.
42. Åstrand, P., Brock, J. C. and Cunnane, T. C., *J. Physiol.*, 1988, **401**, 657–670.
43. Lavidis, N. A. and Bennett, M. R., *J. Physiol.*, 1992, **454**, 9–26.
44. Manchanda, R., D. Phil. Thesis, University of Oxford, 1989.
45. Bennett, M. R., Gibson, W. G. and Poznanski, R. R., *Philos. Trans. R. Soc. London*, 1993, **B342**, 89–99.
46. Stjärne, L., Åstrand, P., Bao, J-X., Gonon, F., Msghina, M. and Stjärne, E., in *Molecular and Cellular Mechanisms of Neurotransmitter Release* (eds Stjärne, L., Greengard, P., Grillner, S., Hökfelt, T. and Ottoson, D.), Raven Press, New York, 1994, pp. 461–496.
47. Hirst, G. D. S. and Neild, T. O., *J. Physiol.*, 1978, **280**, 87–104.
48. Hirst, G. D. S. and Neild, T. O., *J. Physiol.*, 1980, **303**, 43–60.
49. Finkel, A. S., Hirst, G. D. S. and van Helden, D. F., *J. Physiol.*, 1984, **351**, 87–98.
50. Hirst, G. D. S., De Gleria, S. and van Helden, D. F., *Experientia*, 1985, **41**, 874–879.
51. Cassell, J. F., McLachlan, E. and Sittiracha, T., *J. Physiol.*, 1988, **345**, 31–49.
52. Jobling, P. and McLachlan, E. M., *J. Physiol.*, 1992, **454**, 83–105.
53. Manchanda, R., *Electro and Magneto Biol.*, 1994, **13**, 99–106.
54. Huizinga, J. D., Liu, L. W. C., Blennerhassett, M. G., Thunerberg, L. and Molleman, A., *Experientia*, 1992, **48**, 932–941.
55. Garfield, R. E., Thilander, G. and Blennerhassett, M. G., *Can. J. Physiol. Pharmacol.*, 1992, **70**, 481–490.

Received 31 January 1995; revised accepted 21 April 1995

Flexicrystallography: Curved surfaces in chemical structures

Alan L. Mackay

Department of Crystallography, Birkbeck College, University of London, Malet Street, London WC1E 7HX, UK

The various periodic minimal surfaces and nodal surfaces, some of which can be related to infinite regular and semiregular polyhedra, extend the vocabulary of crystallography from finite-coordination polyhedra by making visible two-dimensional manifolds in which atoms may be seen to reside. Recently developed computer programs have greatly facilitated advances in this area. Traditional crystallographic geometry does not apply to curved structures but may be generalized to finite and infinite arrangements, where strict equivalence is replaced by quasi-equivalence.

Especially in biology, assemblies of many copies of the same unit, glibly characterized as imperfect crystals, can often be seen to have regularities outside the compass of the 230 space groups.

The term *'flexicrystallography'* is suggested for the geometric regularities observed in what de Gennes calls 'soft matter'. Vesicles, myelin, plant prolamellar bodies and zeolites provide examples of one-, two- and three-dimensional flexicrystals.

In 1934, J. D. Bernal and Dorothy Crowfoot took an X-ray diffraction picture of a crystal of pepsin which showed that it had a definite structure down to atomic resolution. From 1927, W. T. Astbury had been obtaining X-ray diffraction pictures, full of potential information, from silk, wool, cotton and other fibres and the new world of molecular biology (Astbury's words) had been discovered. Bernal and Astbury made a mutual agreement, like the Spaniards and the Portuguese dividing the New World of the Americas between themselves by the Treaty of Tordesillas in 1494, that Bernal should take the globular proteins and Astbury the fibrous for their research.

The methods of X-ray crystal structure analysis since then have proved immensely successful, some 200,000 structures now being known. They require crystals and the price paid for this kind of microscopy is the assumption that all the unit cells in the crystal are identical. Crystallization normally ensures this, but most natural structures are not crystals and their analysis has been left relatively behind, although, for example, TMV (tobacco mosaic virus) particles have now been successfully examined to atomic resolution from ordered textures. Quasicrystals and incommensurate modulation have also required the extension of crystallographic notation and so have fibres and helices, where structural units are basically linear and thus one-dimensional. We now examine structures where atoms or molecules are arranged in curved surfaces and finite and infinite two-dimensional manifolds are prominent.

In 1967, Bernal, influenced by D'Arcy Thompson's 'Growth and Form' and perhaps by Ernst Haeckel's call for a 'crystallography of organic forms', elaborated his earlier ideas for a 'generalized crystallography', which would systematize those structures, often made of only a few types of molecule, which were ordered but not crystalline. The Theoretical Biology Club in Cambridge of the 1930s, which included J. D. Bernal, C. H. Waddington and J. Needham, were convinced that liquid crystal structures were the bridge between living and nonliving matter. The concept of hierarchic levels of organization (integrative levels) was dominant and was much emphasized by Bernal, Needham and by Whyte[1]. However, liquid crystals, vesicles and fibres usually do not fit into the formalism of the 230 space groups. Five-fold symmetry also is outside this framework, although in 1931 Hermann pointed out that textures might have icosahedral symmetry. Combination of this idea with that of hierarchy led to the prediction of quasicrystals, which were discovered in 1984. Pauling's demonstration of the α-helix with 3.6 residues per turn[2], incommensurable with the screw axes of the space groups, showed that crystallography was not enough. Recent developments have advanced towards what we might call non-conformist crystallography or flexicrystallography, where cognizance is taken of structures outside the framework of orthodox crystallography. The key concept is that of the curvature of one-, two- and three-dimensional manifolds. A wide range of curved two-dimensional manifolds is now available as mathematical objects which can be used, as the 230 classical space-groups are used, for conceptualizing the arrangements of atoms and molecules in space. The most important set of these are the periodic minimal surfaces (of which now about 50 are known).

Since the 230 space groups supply the complete answer to the requirement of positioning identical units in three dimensions, the concept of exact equivalence has to be replaced by that of quasiequivalence, as proposed formally by Caspar and Klug in 1962 in connection with the arrangement of more than 60 virus particles on the surface of a sphere. Problems of the quantification of symmetry thus arise since exact identity is ruled out.

Curvature of a surface

At each point (apart from singularities) a surface has two principal curvatures, k_1 and k_2, in perpendicular planes, each including the surface normal. $1/k_1 = r_1$ and $1/k_2 = r_2$, the radii of the tangent circles. The Gaussian curvature K is then k_1k_2 (dimensions L^{-2}) and the mean curvature H is $(k_1 + k_2)/2$ (dimensions L^{-1}). $2H = \text{div } n$, the divergence of the unit vector normal to the surface.

A soap film with equal pressures on its two sides has $H = 0$ (or, if the pressures are different, H is proportional to the pressure difference) and is called a minimal surface. Any small closed loop on such a surface has a minimum area. Given a boundary, a surface across it has $H = 0$, and there may be several, may be stable or unstable (with respect to small displacements).

On a curved surface of radius R the area of a spherical triangle, bounded by arcs of great circles, is given by the spherical excess and is $A = R^2(\pi - A - B - C)$, where A, B and C are the internal angles of the spherical triangle. Thus, there are no similar triangles in spherical and hyperbolic trigonometry, no parallel translations and, consequently, no crystallography.

On a curved surface the perimeter of a small circle of radius r is $2\pi r(1 - Kr^3/3 + O(r^5))$.

The Gauss–Bonnet theorem, which follows from Stokes' theorem, states that the Gaussian curvature integrated over the surface is equal to the integral of the angles between successive segment of the boundary if the latter is made up of geodesics.

Curved surfaces such as those appearing in membranes may be finite or infinite. Finite vesicles, surfaces of lipid molecules surrounding the water phase, may, for example, be spherical or toroidal. Fullerenes[3] may be spherical or tubular or perhaps even infinite. Lyotropic liquid crystal phases may be infinite and periodic. The discovery and classification of periodic curved minimal surfaces has recently advanced rapidly.

Periodic minimal surfaces

Periodic structures have to belong to one or the other of the 230 space groups of classical crystallography and it was the achievement first of H. A. Schwarz and G. R. Neovius, then of A. H. Schoen (1961), then of S. Hyde and S. Andersson and finally of W. Fischer and E. Koch, to find how patches of curved minimal surface can be repeated to give infinite, three-dimensionally periodic, minimal surfaces. In concept, a patch of soap film is hung across the asymmetric region of a particular space group and is repeated by the symmetry operations of that group to give a continuous surface dividing all space into two separate regions (which may be congruent or different). The film where it is repeated by mirror planes and diad axes must join smoothly so that H is zero everywhere. We have here a compromise between locally non-Euclidean sheets and long-range periodicity.

When H is zero, the two curvatures are equal and opposite and their product K, the Gaussian curvature, is everywhere nonpositive. K may be zero at certain flat points or umbilics, for example, where a threefold axis intersects the surface. Since the Gaussian curvature is negative almost everywhere, the two-dimensional surface resembles locally the hyperbolic plane and has a non-Euclidean metric. The perimeter of a circle is greater than $2\pi r$ and thus the hexagonal close packing of spheres, appropriate for the plane, no longer fits tightly. In the plane the mean number of neighbours, the coordination number N, in a triangulated arrangement of points, is 6. On a surface with K negative, N must be greater than 6 and with K positive (as on a sphere), N must be less than 6. Indeed, we know that, as in fullerenes, in order to make a closed shell of hexagons, 12 disclinations where we have rings of five are necessary (or their equivalent in squares and triangles). If we insist on a few rings of seven in a hexagonal net, then a tangle like seaweed may result. A periodic minimal surface is an ordered structure which answers to this requirement for $N > 6$. In anticipation such, as yet hypothetical, structures of carbon have been provisionally named *schwarzites* (by V. Elser after H. A. Schwarz).

Patches, with a locally non-Euclidean metric can thus be sometimes reconciled with three-dimensional crystallinity.

Exact minimal surfaces are quite hard to compute. Exact methods involve the Weierstrass integrals (which are line integrals in the complex plane with elliptic functions). Some of the Weierstrass integrals have recently been solved as algebraic functions which are directly available in modern mathematical computer systems such as Mathematica. Finite element methods are also effective, and have now been brought into a usable form in the Surface Evolver program of K. Brakke, freely available through the Minnesota Geometry Centre for the working chemist who may wish to conjure up a surface to compare it with, for example, an actual zeolite or lipid structure.

The problem of plotting soap-film and other surfaces

Brakke's 'Surface Evolver'

At the Minnesota Geometry Centre, K. Brakke has developed a program Surface Evolver (freely available by electronic mail) which makes the finite element methods of engineering available for problems like those of soap films of arbitrary shape. Many more elaborate calculations involving contact angle, gravity, other functions of curvature, etc., are possible. It works extraordinarily fast on ordinary 486 PC and other computers and greatly extends the possibilities available to individual chemists.

Mathematica

The mathematical program system Mathematica, developed by Steven Wolfram, enables many algebraic, numerical and graphical calculations to be carried out automatically. In particular, we have used the operation ContourPlot3D, which expeditiously draws the surface $f(x, y, z) = 0$, even though the variables may not be separable. Many of the figures here used this. Hypergeometric functions, earlier available only in obscure tabulation, can be called directly for plotting surfaces. The importance of this development on the outlook of chemists towards mathematics cannot be overemphasized. We are in a new period.

has become simplified by the revival of interest in structure factor graphs[4]. These are nodal surfaces in the various crystallographic space groups. When calculations of scattering factor were done by hand, to see if they matched the X-ray diffraction data and thus confirmed the postulated structure (or not), it was useful to see the effect on the calculated structure factors of moving a particular atom and its symmetry-related equivalents. The structure factor in question, a sum of sine waves, could be plotted in space, but, in fact, Bragg and Lipson[5] used only sections. The nodal surface of the structure factor in question divides space into two regions; on one side of the nodal surface the atom contributes positively, on the other side it contributes negatively and on the surface it has zero contribution to the diffracted intensity. The position of an atom can be adjusted accordingly. Now with computer graphics these surfaces can be readily plotted in three dimensions.

It turns out that these nodal surfaces are sometimes close to minimal surfaces with the same space group. Here the symmetry-related sine waves corresponding to a form $\{h, k, l\}$ for a particular space group are added. To get a still better approximation it may be useful to add structure factor expressions for two or more forms. This method of expansion of an arbitrary surface as a series of Fourier terms had been used to describe Fermi surfaces, some of which are similar to the standard periodic minimal surfaces.

The infinite polyhedra (skew polyhedra), found by Coxeter and Petrie when they were undergraduates, are limiting cases where the curved surfaces become flat panels. These have become familiar through their occurrence in zeolite frameworks.

Electrostatic equipotentials

A further important class of surfaces arise in crystallography while considering the Coulomb forces in ionic structures. The zero equipotential between the positive and negative ions in the cubic CsCl (which is primitive cubic with one ion at (0, 0, 0) and the other at $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$) is close to the P-surface (the simplest periodic minimal surface). This in turn is close to the nodal surface corresponding to the {1, 0, 0} reflections (three equivalent sine waves). Additional terms can be added to give any required degree of accuracy[6]. Calculation of the potential in a lattice of charges and the evaluation of the Madelung constant are related to the very slow convergence of the lattice sum. However, Ewald's original observation[7] that the potential involved elliptic integrals provides a powerful method which can be used for calculating equipotential surfaces.

Plotting nodal surfaces

Given, for example, the Mathematica program system and an additional program Contour Plot3D, this and other nodal surfaces can be directly generated. However, pictures of the various nodal surfaces can be obtained faster by stochastic method. Either a point can be followed as it runs over the surface following a geodesic, or particle may be reflected back and forth inside one of the two regions of the system and its position can be marked whenever it strikes a surface. This method gives a random dot pattern where the surface is evident only if a stereo pair is constructed. Such pairs could serve as a test for stereopsis after the method developed

Nodal surfaces

Nodal surfaces in a standing wave field are the places where the amplitude is zero. They became familiar in Chladni's demonstration of the movement of fine sand on vibrating plates, although, of course, the Pythagoreans knew of the one-dimensional case of the monochord. Where the amplitude is zero the phase angle is indeterminate.

The physical (as well as the mathematical) formulation of material periodic surfaces by the use of nodal surfaces in acoustic, optical and perhaps X-ray standing-wave fields is also advancing.

The contribution to the reflection h, k, l of one atom with scattering factor f and fractional coordinates x, y, z with respect to a predetermined origin in a crystal with a particular space group crystal is given by

$$F_{hkl} = f(\cos[2\pi(hx + ky + lz)] + i \sin[2\pi(hx + ky + lz)]).$$

The number and phases of the symmetry-equivalent atoms depend on the space group. There may be up to 192 atoms for the most symmetrical cubic case. A low-symmetry space group may generate no surfaces. This was the summation used by Bragg and Lipson.

A variant, relevant for building up arbitrary surfaces, consists in adding all symmetry-equivalent reflections of the form $\{h, k, l\}$ to give an image. (When there are no translation elements the two types of sum may be the same.) The sum may usually be expressed more concisely preferring the products of sine and cosine terms required by the early form of the fast Fourier transform used in the hand summation method of the Beevers–Lipson strips extant before the advent of the electronic computer.

Tables of such structure factor and electron density expressions were developed by K. Lonsdale and W. F. Astbury and are now embedded in the International Tables for Crystallography.

The nodal surface is the plot of the values of (x, y, z) for which $f_{x,y,z}$ is zero. (Both real and imaginary parts of f must be zero.)

A table of nodal surfaces equivalent to various periodic minimal surfaces has been given by von Schnering and Nesper[4].

by Bela Julesz. The diffusion of atoms in a molecular sieve can be simulated and parameters such as the Lyapunov coefficient calculated.

Zeolites

Many of the frameworks of zeolite minerals can be seen to lie on periodic minimal surfaces. The clearest example is that of faujasite, where large tetrahedral units – four tubes meeting at 109° to each other, follows the D-surface with a diamond type of structure.

Extension to N-dimensions

Since the expression $\cos x + \cos y + \cos z = C$ gives, as C varies, a set of surfaces parallel to the $\{1\,0\,0\}$ nodal surface, this could be embodied into a generalized form $\cos x + \cos y + \cos z + \cos w = 0$, which would be a four-dimensional nodal boundary. The whole set of parallel surfaces would be obtained as sections through this four-dimensional object. The curved three-dimensional space S^3, which is the three-dimensional analogue of the surface of a sphere, has been used, particularly by Sadoc and Kléman in Orsay, for modelling amorphous materials.

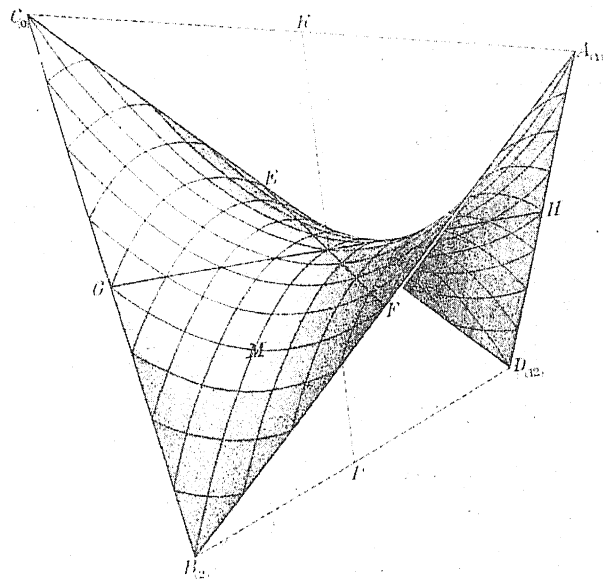
Finite surfaces

Vesicles are small micelles of lipid molecules containing a water phase. Usually, they are spherical, but, recently, toroids[8,9] and elaborate helically twisted cylinders[10], have been observed. They are the most organized and perhaps tractable of the great variety of lyotropic structures encountered in biology.

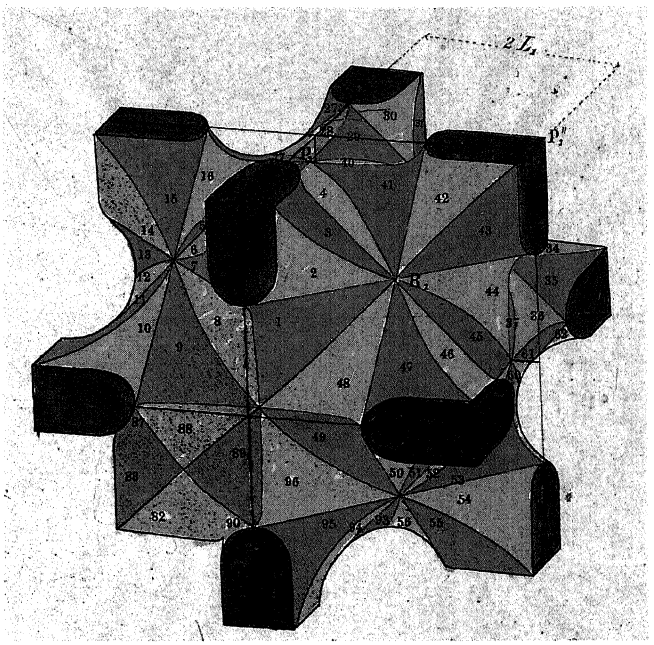
A closed surface where, for a given volume, the integral of H^2 over the surface is a minimum is called a Willmore surface. For a symmetrical torus this minimum is achieved for a ratio of $\sqrt{2}$ between the two radii. It has also been shown that this integral is invariant with respect to a conformal transformation of the torus into a Dupin cycloid[9].

Curiously, as pointed out also by Andersson[11], the finite crystal forms $\{h, k, l\}$ of traditional mineralogy can be plotted by using expressions similar to the nodal expressions given above for periodic minimal surfaces but replacing sine and cosine by the hyperbolic sine and cosine. Thus, for example with $\{1, 0, 0\}$, $\cosh x + \cosh y + \cosh z = 1000$ gives a cube, and with $\{1, 1, 1\}$, $\cosh(x + y + z) + \sinh(x + y + z) = 1000$ gives a tetrahedron. In the complex plane cosine and cosh are at right angles to each other.

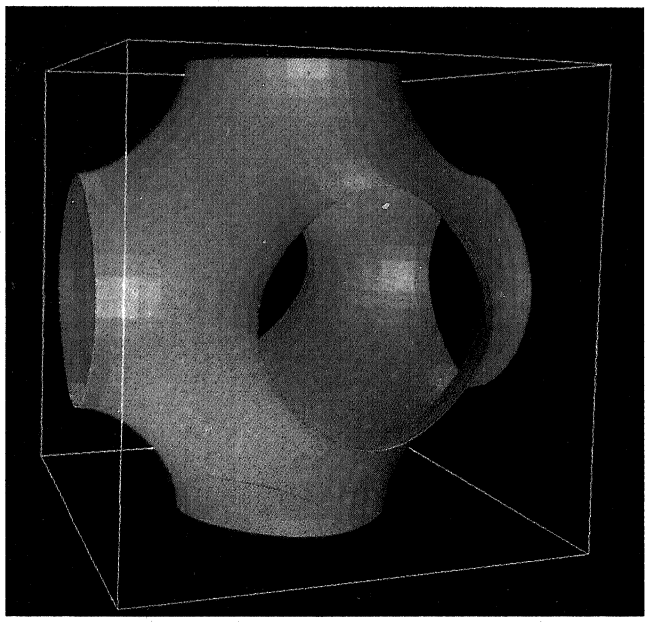
1



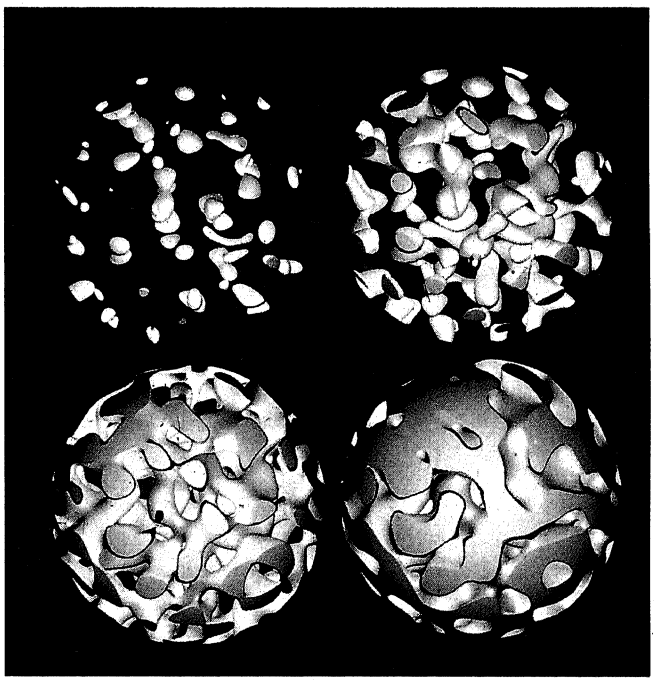
2



3

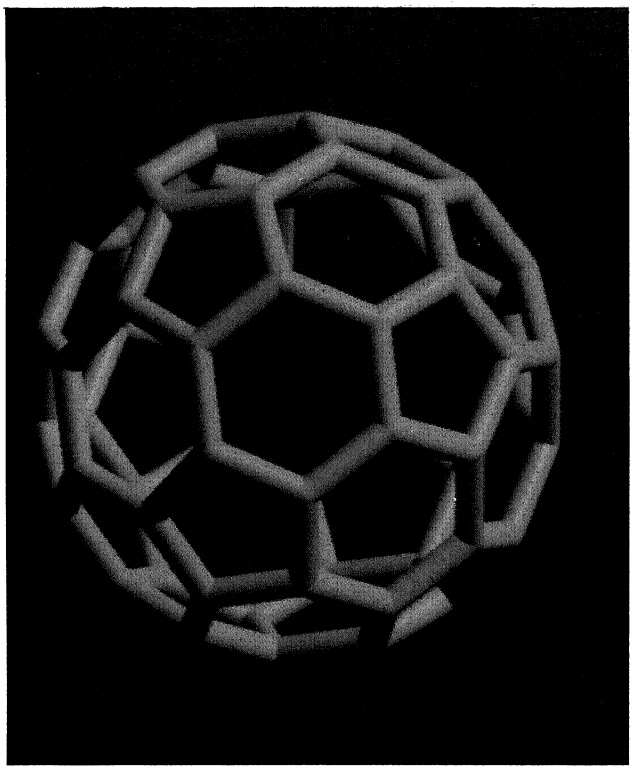


4

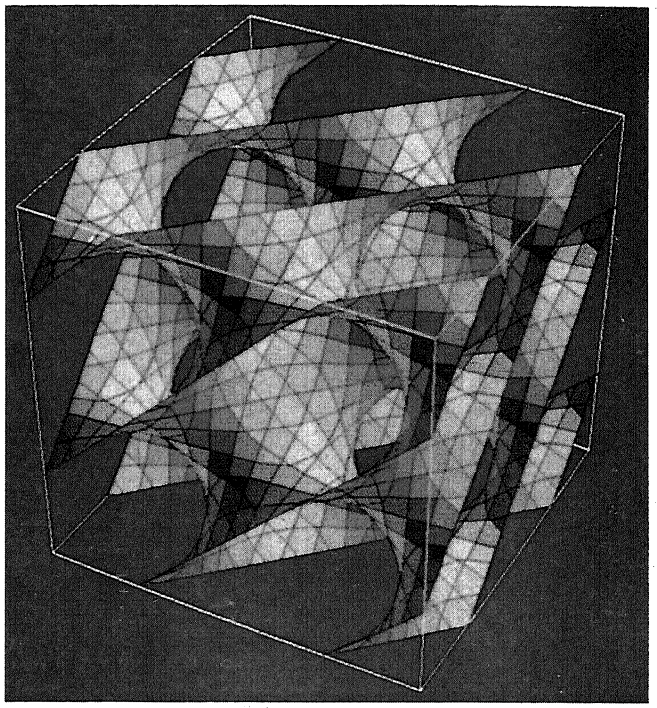


For captions, see page 160.

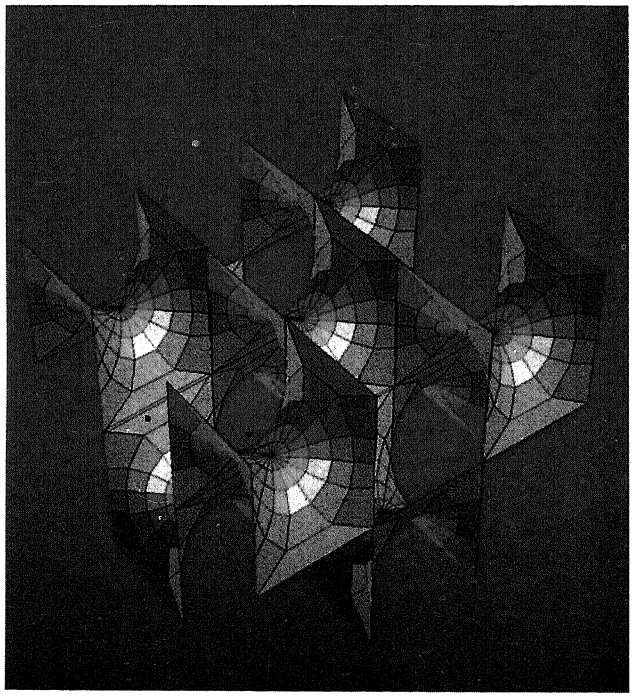
5



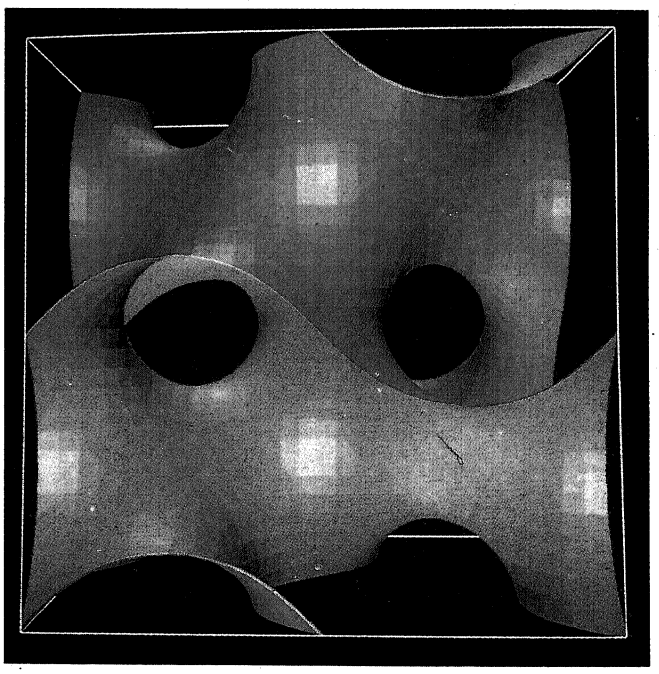
6



7

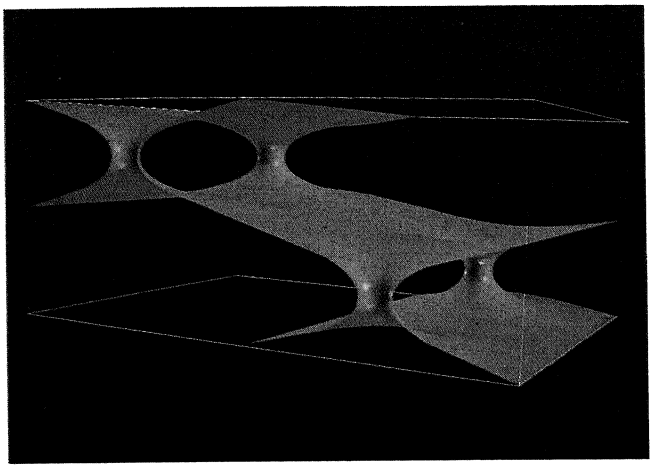


8

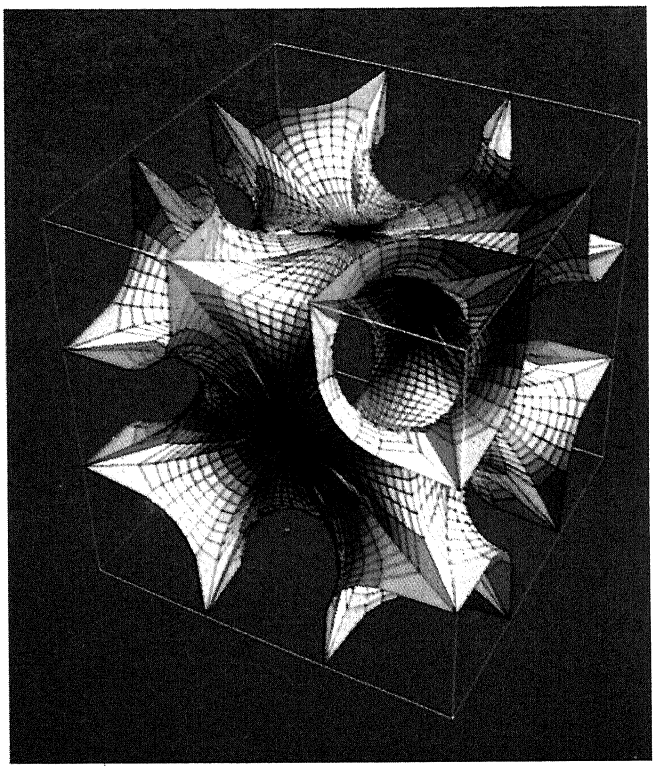


For captions, see page 160.

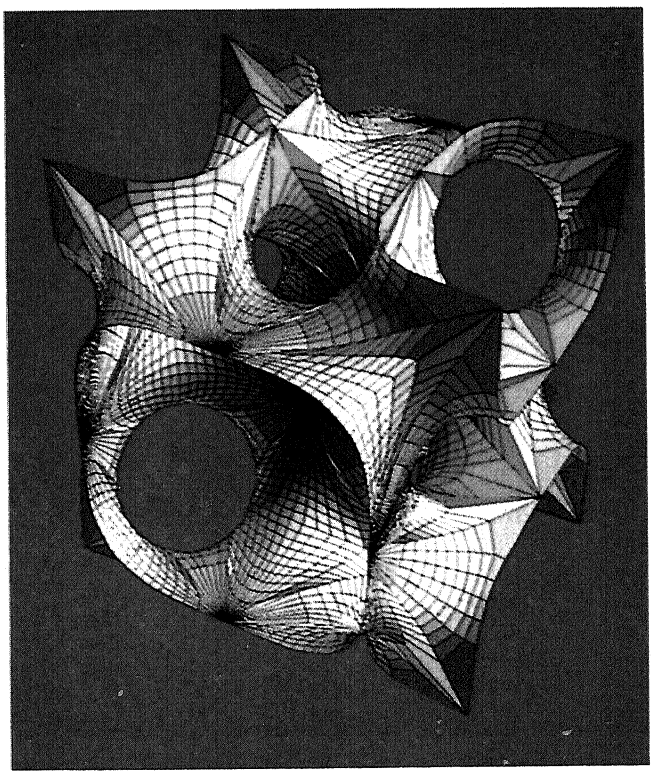
9



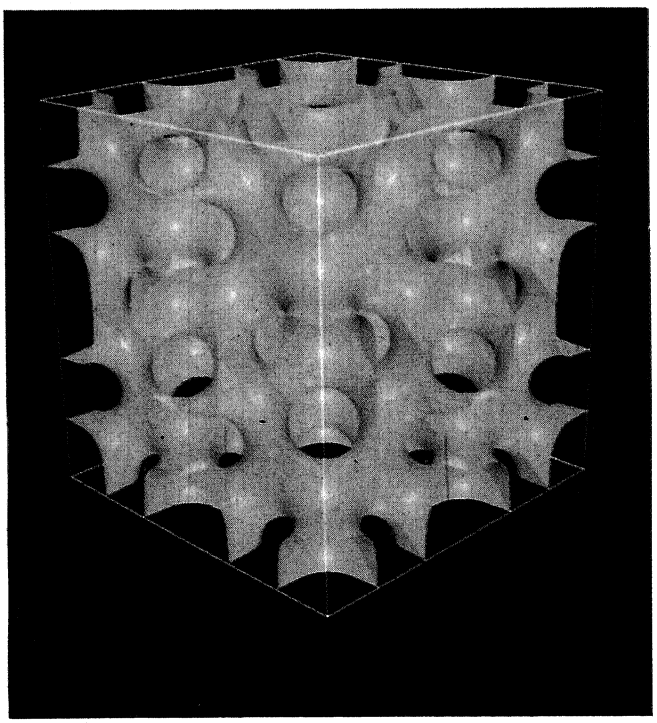
10



11

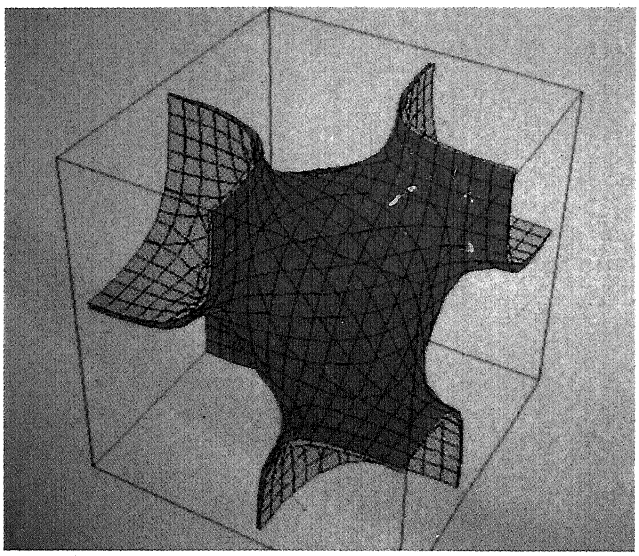


12

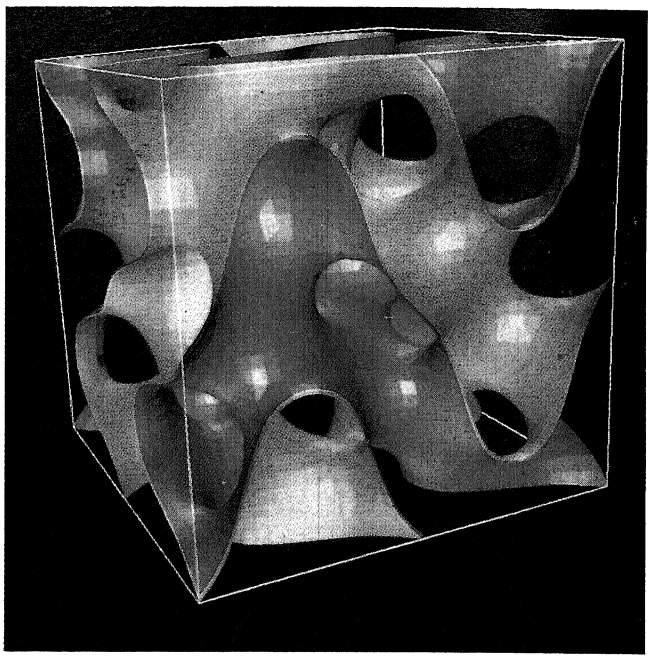


For captions, see page 160.

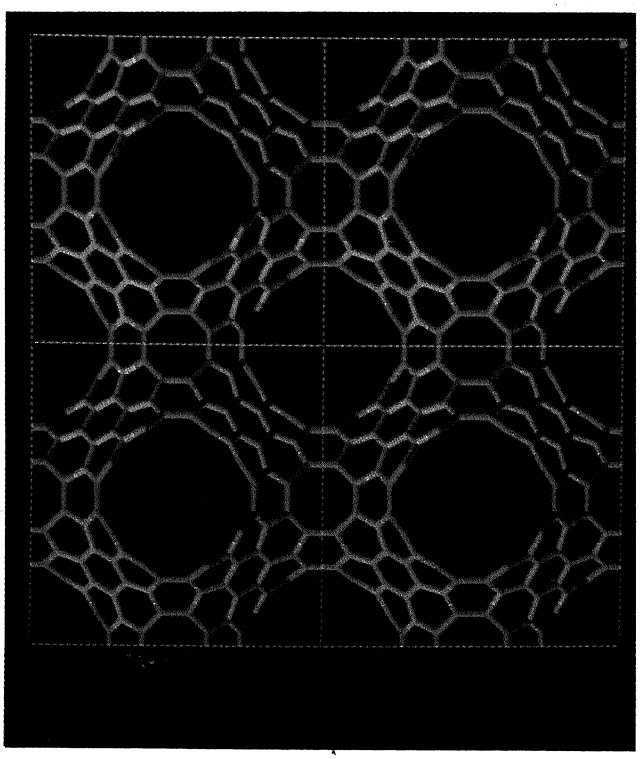
13



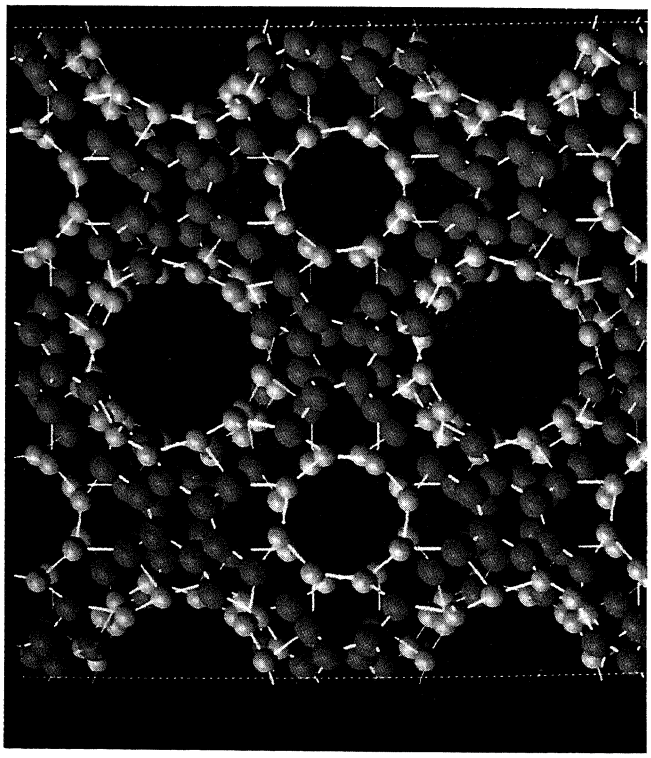
14



15

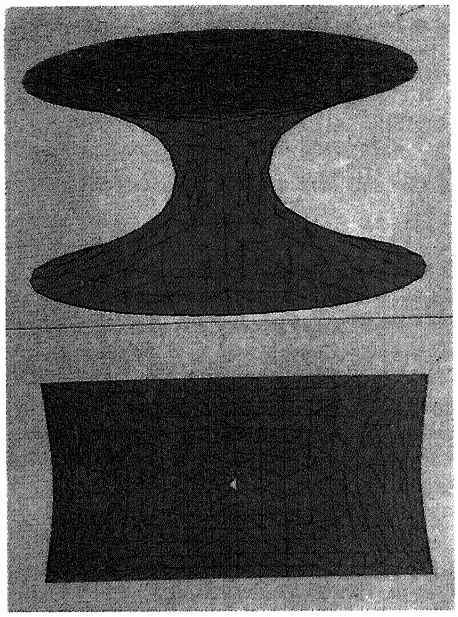


16

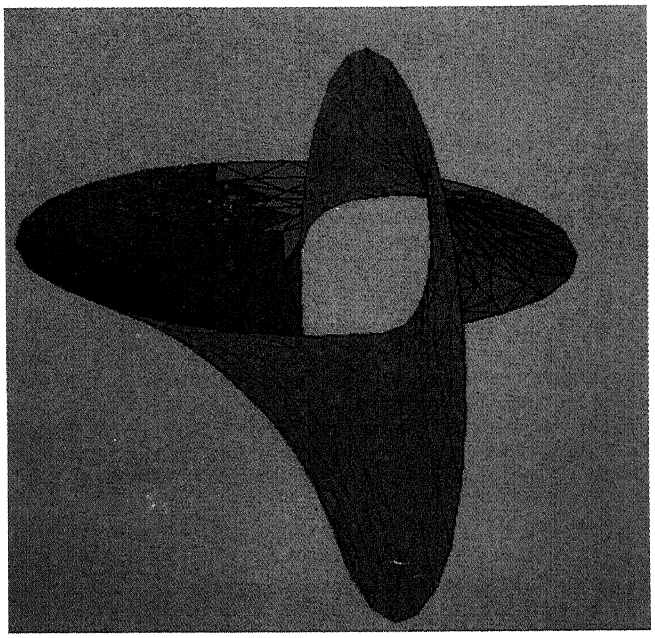


For captions, see page 160.

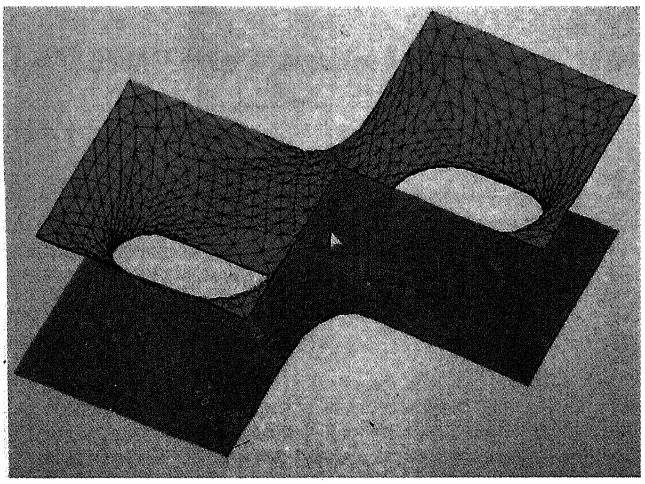
17



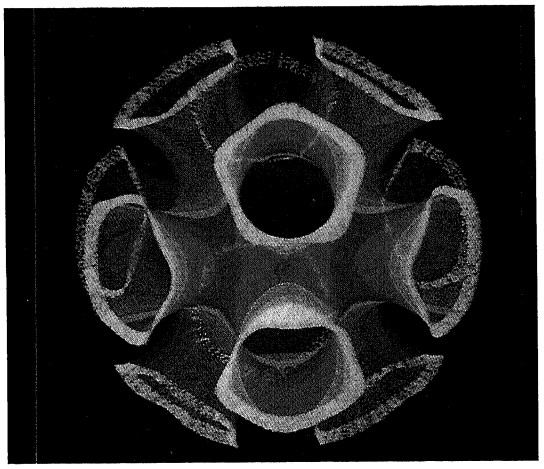
18



19



20



For captions, see page 160.

Figure captions

Figure 1. A soap film surface over a tetrahedral frame. Figure from H. A. Schwarz (1865/1890).

Figure 2. The Neovius surface (figure from Neovius' thesis, 1883). One compartment is connected by 12 tubes in the [1 1 0] directions to the next. The space group is $Im\bar{3}m$ if the two sides of the surface are taken to be identical. The 96 equivalent curved unit triangles have angles 45°, 45° and 30° with a deficit of 60° per triangle.

Figure 3. The P-surface (generated as the surface $\cos x + \cos y + \cos z = 0$). This was one of the five found by Schwarz (1865–). The space group is $Pm\bar{3}m$ (thence P for *primitive*) if the two sides of the surface are different, and $Im\bar{3}m$ if they are the same. There are six tubes in the [1 0 0] directions. This surface is also close to that of the $V = 0$ equipotential in the structure of cesium chloride (where Cs^+ is at 0, 0, 0 and Cl^- at $\frac{1}{2}, \frac{1}{2}, \frac{1}{2}$). The 48 equivalent asymmetric triangles have angles 90°, 45°, 30°.

Figure 4. A statistical nodal surface generated by R. Phippen, after J. W. Cahn, by adding 20 sine waves all of the same wavelength but with random phases, directions and amplitudes. The surface is contoured at 0.5, 0.375, 0.25 and 0.125 and may represent an oil/water emulsion.

Figure 5. C_{60}, the molecule of buckminsterfullerene, where 60 atoms are distributed on the surface of sphere, each atom being identically situated. There are two different bond lengths, so that the figure is not exactly that of semiregular Archimedean solid, the truncated icosahedron. Each hexagon has six neighbours and each pentagon five, so that the mean number is less than six.

Figure 6. The D-surface (D for *diamond*) (due to Schwarz). Forty-eight of the tetrahedral saddles shown in Figure 1 are contained in a face-centred cubic cell (space group). If the tetrahedral bonds in the diamond structure are inflated to tubes, so that the volume inside the tubes is equal to that outside, then this D-surface is obtained. Alternatively, the D-surface can be seen as the surface separating the two bonding networks in the structure of cuprous oxide.

Figure 7. The T-surface (T for *tetrahedral*) found by Schwarz. The *c/a* ratio may be varied but the surface is drawn here with $c/a = 1$.

Figure 8. The G-surface (G for *gyroid*) is a remarkable surface found by Alan Schoen around 1960. The space group is $Ia\bar{3}d$ if both sides of the surface are considered equivalent. The two subspaces are mirror images of each other. It is, in a sense, midway between the P- and D-surfaces.

Figure 9. The o-CLP-surface (CLP stands for *crossed layers of parallels in Schoen's nomenclature*, o for *orthorhombic*). One of the 13 new surfaces found by Schoen. It has variable cell dimensions. Drawn with these proportions it shows the way in which catenoidal bridges may connect almost parallel sheets as perhaps in the transformation of a layer-structured lipid to cubically cross-connected structure.

Figure 10. The I-WP-surface (IWP is I (body-centred) *wrapped package* in Schoen's nomenclature) with the space group $Im\bar{3}m$. Here the two subspaces are no longer equivalent but have a volume ratio of $\sqrt{3}$ to 2. Cvijović has given an exact expression for the surface in terms of hypergeometric functions (used here for plotting it). There are eight tubes in the [1 1 1] directions.

Figure 11. The same IWP cell plotted with a different origin.

Figure 12. A $2 \times 2 \times 2$ block of the IWP surface showing how an extended volume builds up. It might be used as a heat exchanger, or for separating two chemicals for a binary reaction, etc.

Figure 13. One of the eight tetrahedral joints which make up the unit cell of the IWP surface. This was plotted by Mathematica with the ContourPlot3D program as the nodal surface of the superposition of several sine waves.

Figure 14. The S-surface, plotted as the nodal surface of the form {2 1 1} with the function $\cos 2x \sin y \cos z + \cos x \cos 2y \sin z + \sin x \cos y \cos 2z = 0$. The space group is $Ia\bar{3}d$ if the two sides are considered identical.

Figure 15. The P-surface in which a graphite net has been embedded with the introduction of disclinations making some rings of six carbon atoms into nonplanar rings of eight. Calculations indicate that this negatively curved structure should be more stable than that of spherical (positively curved) buckminsterfullerene but it remains hypothetical.

Figure 16. A graphite net mapped on to the G-surface (and on to a number of the other surfaces).

Figure 17. The surface of zero mean curvature between two parallel rings is a catenoid, but there are two solutions; one stable (below) and one unstable (above). The stable surfaces are very easily mapped using the finite analysis methods provided by Surface Evolver. Unstable surfaces are more difficult to model and require additional conditions (such as a fixed internal volume).

Figure 18. By changing the positions of the rings in the previous example, the area of the surface of a soap film hanging between two perpendicular circles, each through the centre to the other, can very rapidly be found numerically with the Surface Evolver program. Algebraic treatment is reported to be so far impossible.

Figure 19. Other difficult problems, such as finding the surface of two joined square catenoids, which appear in certain surfaces discovered by Fischer and Koch, are readily solved by the Surface Evolver.

Figure 20. By following geodesic lines on a surface, in this case the dodecahedral nodal surface formed by adding six sine waves, the surface can be mapped out. The trajectory of a point is deterministically chaotic and eventually covers the whole surface.

Illustrations

The plates illustrate a number of shapes now being produced. They are explained in the associated captions.

1. Whyte, L. L., Wilson, A. G. and Wilson, D., *Hierarchical Structures*, Elsevier, New York, 1969.
2. Pauling, L. and Corey, R. B., *J. Am. Chem. Soc.*, 1950, **72**, 5349.
3. Kroto, H. W. and Walton, D. R. M., *The Fullerenes*, Cambridge University Press, 1993.
4. von Schnering, H. G. and Nesper, R., *Z. Phys.*, 1991, **B83**, 407–412.
5. Bragg, W. L. and Lipson, H. S., *Z. Kristallogr.*, 1936, **95**, 323–337.
6. Mackay, A. L., *Chem. Phys. Lett.*, 1994, **221**, 317–321.
7. Ewald, P. P., *Ann. Phys.*, 1921, **IV64**, 253–287.
8. Fourcade, B., Mutz, M. and Bensimon, D., *Phys. Rev. Lett.*, 1992, **68**, 2551–2554.
9. Fourcade, B., *J. Phys. II France*, 1992, **2**, 1705–1724.
10. Köning, J., Boettcher, C., Winkler, H., Zeitler, E., Talmon, Y. and Fuhrhop, J-H., *J. Am. Chem. Soc.*, 1993, **115**, 693–700.
11. Andersson, S., Jacob, M. and Lidin, S., *Z. Kristallogr.*, 1995, **210**, 3–4.

ACKNOWLEDGEMENTS. I am most grateful to my colleagues, J. Klinowski, D. Cvijović and H. Terrones for help and for providing a number of pictures. Financial help for computing from the Royal Society and University of Paris–Sud has been most valuable and the hospitality of the Indian Institute of Science is particularly acknowledged.

Received 12 June 1995; accepted 13 June 1995

Observed and theoretical acceleration response spectra in the Tehri region: Implications for seismic hazard in the region

K. N. Khattri

Wadia Institute of Himalayan Geology, 33, General Mahadeo Singh Road, Dehra Dun 248 001, India

This paper presents an estimate of the seismic hazard in the Tehri region of the Himalaya. Theoretical simulations of strong motion expected in the vicinity of the fault of a $M_s = 7$ earthquake are used in conjunction with observed strong motion data for the 1991 Uttarkashi earthquake to conclude that for such an earthquake the peak acceleration at Tehri can be ~ 0.8 *g*. The corresponding spectral acceleration at $T = 1$ s can be ~ 0.98 *g*. Analysis of the strain budget of the region and a finite probability of occurrence of a great earthquake ($M_s > 8$) in the next 100 years or so is estimated. The spectral acceleration (T~ 1–2 s) in the near-field region of such an earthquake is estimated to be of the order of 1 *g*.

ESTIMATION of seismic hazard is ordinarily based on one-parameter characterization of ground shaking, such as peak ground acceleration, peak ground velocity or peak ground displacement. The maps specify values of the chosen parameter at various sites with a given probability of exceedance in a chosen time span. Such single-parameter descriptions are most appropriate for characterizing the damage potential to short-period structures. However, in assessing the damage potential to complex structures such as dams and power plants, where a range of periods of the main structure and the substructures are involved, the details of the strong ground motion histories are needed to estimate properly the seismic hazard. For the purposes of seismic hazard characterization, the information in strong-motion time history is transformed into response spectrum, which specifies the levels of expected spectral accelerations at various periods.

Modelling is a standard scientific technique used for identifying the essential features of natural phenomena. Models are used to extrapolate and predict phenomena in domains for which empirical observations may not be available, but are needed. This is especially true in strong-motion seismology, as relatively few strong-motion accelerograms have been recorded so far, and the expected strong-motion time histories in earthquake-prone zones are needed to determine the parameters for earthquake-resistant design of structures.

In this paper we use the composite source model to represent a hypothetical earthquake similar to the 1991 Uttarkashi earthquake (M_s =7) placed near Tehri, and use wave propagation theory to simulate the strong ground motions at Tehri[1–3]. We compare the results with those recorded for the Uttarkashi earthquake and show that the simulated strong ground motions are realistic.

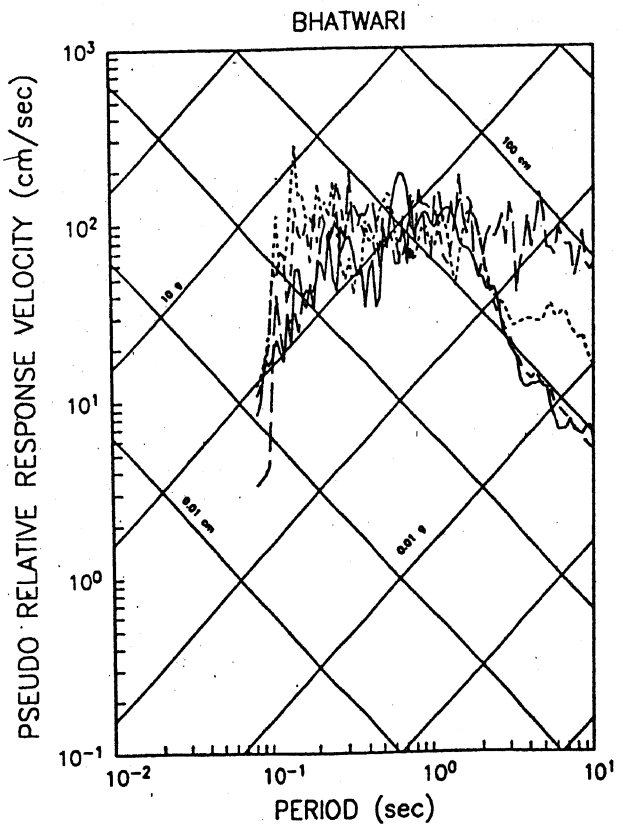


Figure 1. The observed pseudorelative response spectra at Bhatwari for the 1991 Uttarkashi earthquake ($M_s = 7$) are shown by solid line. The synthetic spectra for Uttarkashi earthquake are shown by short dashes (no filtering of the synthetic accelerogram) and by intermediate dashes (filtering of the synthetic accelerograms with a filter equivalent to the one used in the observed accelerogram data processing. The long dashes show the response spectrum for an $M_w = 8.5$ earthquake[2]. Note that the agreement between the observed and the synthetic response spectra is good for periods longer than 0.2 s. The damping in all cases is at 5%.

We compute the response spectra for the simulated event, and compare them with the response spectra obtained at Bhatwari and Uttarkashi for the 1991 Uttarkashi earthquake. These sites were at comparable distances from the surface projection of the respective fault plane edge as was the simulated event from the surface projection of the fault plane edge of the hypothetical earthquake.

The magnitude of the most damaging earthquake that may strike the region in the next 100 years or so is also estimated using the available data on the past history of great earthquakes, the geological model of active tectonics in the region, and the strain rates obtained from a variety of geological and geophysical measurements as well as seismological data. The corresponding peak as well as spectral accelerations expected in the earthquake fault region are estimated. The peak accelerations expected on the basis of probabilistic analysis of seismological data are estimated and compared with those determined by the deterministic methods mentioned above.

The modelling procedure

The ground motions due to an earthquake depend on the position and nature of the earthquake source and the properties of the intervening medium; density, wave velocity and wave attenuation as a function of space coordinates. Knowing these parameters one can use the available theoretical frameworks to estimate the expected strong ground motions at a site.

Since earthquakes involve a complex process of rock fracturing, the procedure consists of representing the earthquake fault as a superposition of a large number of smaller circular subfaults distributed randomly over the earthquake fault plane. This models the heterogeneous nature of the earthquake faulting process. The number of subfaults with a given radius (R) is proportional to $(1/R^D)$, where D is a fractal dimension. All subevents are assumed to have the same stress drop. Each subfault radiates a Brune source pulse[4, 5]. The sum total of moments of all the subfaults is constrained to equal the earthquake moment. The rupture is initiated at the hypocentre and sets off other subfaults as its rupture front expands and touches their centres. The elastodynamic radiation, for an appropriate model of earthquake source, is propagated to the site of observation by using numerically synthesized Green's functions for an appropriate layered earth structure. For $D = 2$, one obtains an ω^{-2} falling off of the spectrum at high frequencies of the elastodynamic radiation from a composite fault as used here. This is consistent with the empirical observations[6].

Results

Khattri *et al.*[2] and Yu *et al.*[3] used the above approach to model the strong ground motion records of the 1991 Uttarkashi earthquake and showed that the technique is able to reproduce the essential stochastic features of the observed data. The corresponding response spectra were also well matched over a wide range of periods. An example of this is shown in Figure 1. The response effects of the local (shallow) site geology, which is not known, is the most likely cause for the observed differences between the observed and the simulated spectral accelerations. Other factors that can contribute to the differences are inaccuracies in the velocity models used, as well as resonance effects of the severe topography in the region.

In this paper we use the above technique to estimate the nature of strong ground motions at Tehri expected for a hypothetical future earthquake of magnitude

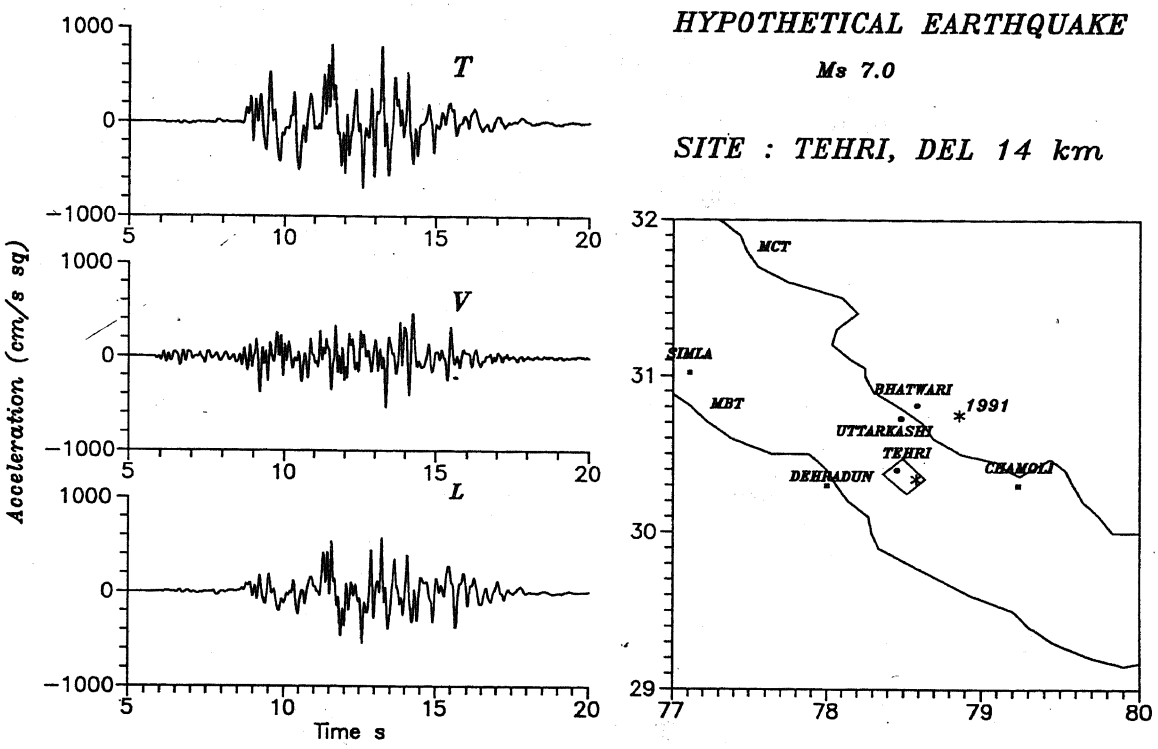


Figure 2. Accelerograms synthesized at Tehri for $M_s = 7$ earthquake. The location of the earthquake is shown as inset in Figure 3.

$M_s = 7$, having similar fault plane solution and focal depth as the 1991 Uttarkashi earthquake. The fault plane of the hypothetical earthquake lies in the vicinity of Tehri, as shown in Figure 2. Tehri is located near the NW corner, but lies within the surface projection of the fault plane. The fault plane is located at a depth of 12 km with a dip of 14° towards NW. The strike of the fault plane is 317°. The area of the model fault plane is 16×24 km^2. This is based on the evidence for the 1991 Uttarkashi earthquake and the empirical relation available between fault dimension and the earthquake magnitude[7]. The slip is of thrust type, with the overthrusted block moving in an approximately SW direction. The velocity model used is shown in Table 1[2, 3]. Note that a top layer of low velocity is included in the model to represent a weathered layer, and (or) sediments. A stress drop of 30 bar is used in the simulation.

The synthetic acceleration time histories at Tehri for the transverse, vertical and longitudinal components are shown in Figure 2. The duration of strong shaking is approximately 8 s. The transverse component shows the largest value of 816 cm/s^2 for the peak acceleration. The peak acceleration values for the longitudinal and the vertical components, respectively, are 573 cm/s^2 and 537 cm/s^2. There is a lower-frequency component in these accelerograms which seems to separate the larger acceleration peaks.

Figure 3 shows the spectra of the acceleration time histories. The spectra show a corner at a frequency of about 0.02–0.04 Hz, which corresponds to a source dimension radius of ~8 km. This is consistent with the subfault size chosen for modelling the earthquake ground motions. At 10 Hz the sudden drop in amplitude is due to the computation of the Green's functions up to that frequency only.

The acceleration response spectra at 5% damping are shown in Figure 4. The highest spectral accelerations of about 2 g are observed in the short period range of 0.1–0.7 s. At a period of 1 s, the spectral acceleration value is ~ 1 g for the T component, ~ 0.7 g for the L component and ~ 0.3 g for the vertical component. At a period of 2 s, the spectral acceleration value is ~ 0.8 g for the L component, 0.55 g for the T component and ~ 0.2 g for the vertical component.

Table 1. Velocity and Q model

Thickness (km)	V_P (km/s)	Q_P	V_S (km/s)	Q_S	Density (g/cc)
0.4	3.5	50	2.00	25	1.80
1.0	5.0	80	2.86	50	2.40
15.0	5.2	4000	2.97	2000	2.60
30.0	6.0	4000	3.43	2000	2.90
Infinity	8.33	1000	4.83	500	3.30

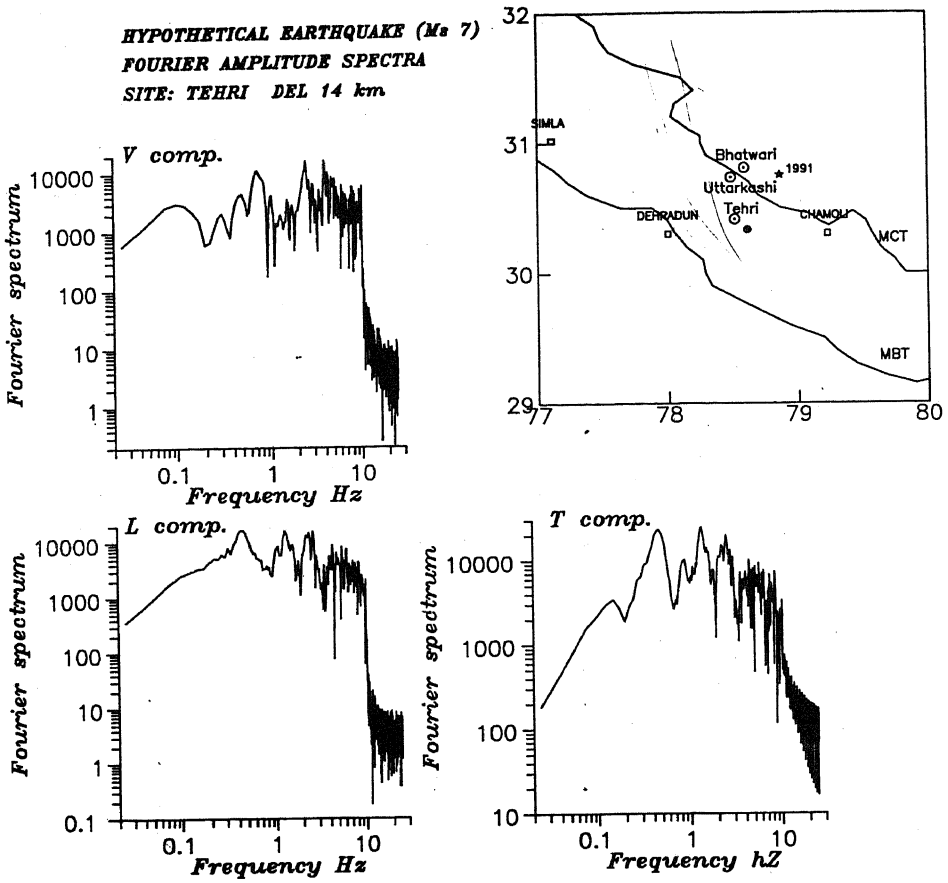


Figure 3. Fourier spectra of accelerograms synthesized for the site at Tehri.

Discussion

Accelerations near the fault zone

We discuss the theoretical results obtained above in the light of the observed strong motions at Uttarkashi and Bhatwari for the 1991 $M_S = 7$ Uttarkashi earthquake. These sites are located close to (within about 2–10 km) the surface projection of the inferred fault plane. Also, they are located in the direction of rupture propagation. In the simulated model, the Tehri site is located over the surface projection of the fault plane, about 7 km inside it. It is also in the direction of rupture propagation.

The observed accelerograms for Bhatwari and Uttarkashi are shown in Figures 5 and 6, respectively. Bhatwari is at an epicentral distance of 25 km and Uttarkashi at 36 km. Although the epicentral distance loses its conventional meaning in the near field of the earthquake, it serves to provide a reference scale in terms of relative positions with respect to the fault plane as well as the fault dimension. The propagation effects will also be governed by this parameter. Peak accelerations of 0.24 g, 0.30 g and 0.19 g have been recorded for L, T and V components, respectively, at Uttarkashi. Similarly, peak accelerations of 0.25 g, 0.24 g, 0.29 g have been recorded for L, T and V components, respectively, at Bhatwari.

The peak acceleration represents the high-frequency component of the energy, and most of it will be contributed by the faulting process near the site. Since both Bhatwari and Uttarkashi are at similar distances from the fault edge, they are likely to have similar values of the peak accelerations. Such is indeed the case, as is documented above. However, at lower frequencies (longer periods) the contributions from the entire fault plane will be important, as they will become increasingly coherent with longer periods (wavelengths). Thus, the entire fault plane will contribute to the energy. The epicentral distance will also play a more significant role in determining the amplitude. This phenomenon can be observed better in the response spectrum.

The response spectra at Bhatwari and Uttarkashi are shown in Figures 7 and 8, respectively. We note that the shorter-period spectral responses (0.1–0.4 s) show comparable levels at the two stations, in keeping with simi-

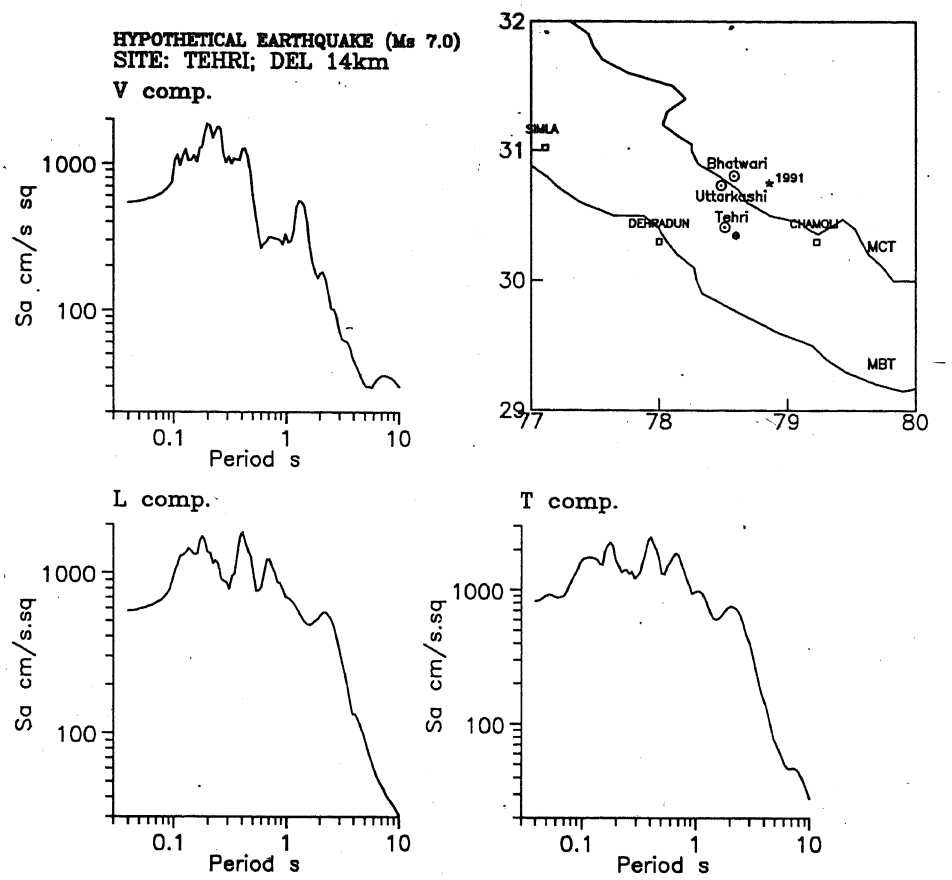


Figure 4. Response spectra at Tehri for 5% damping.

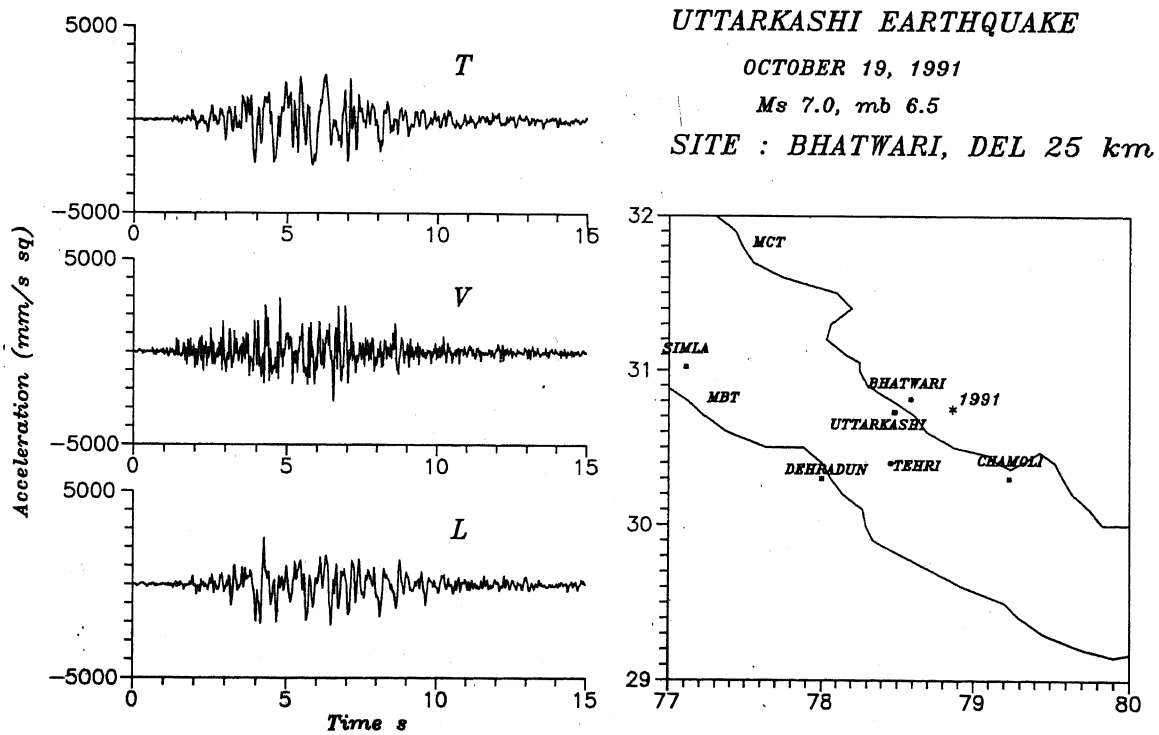


Figure 5. Accelerograms recorded at Bhatwari for the 1991 Uttarkashi earthquake ($M_s = 7$).

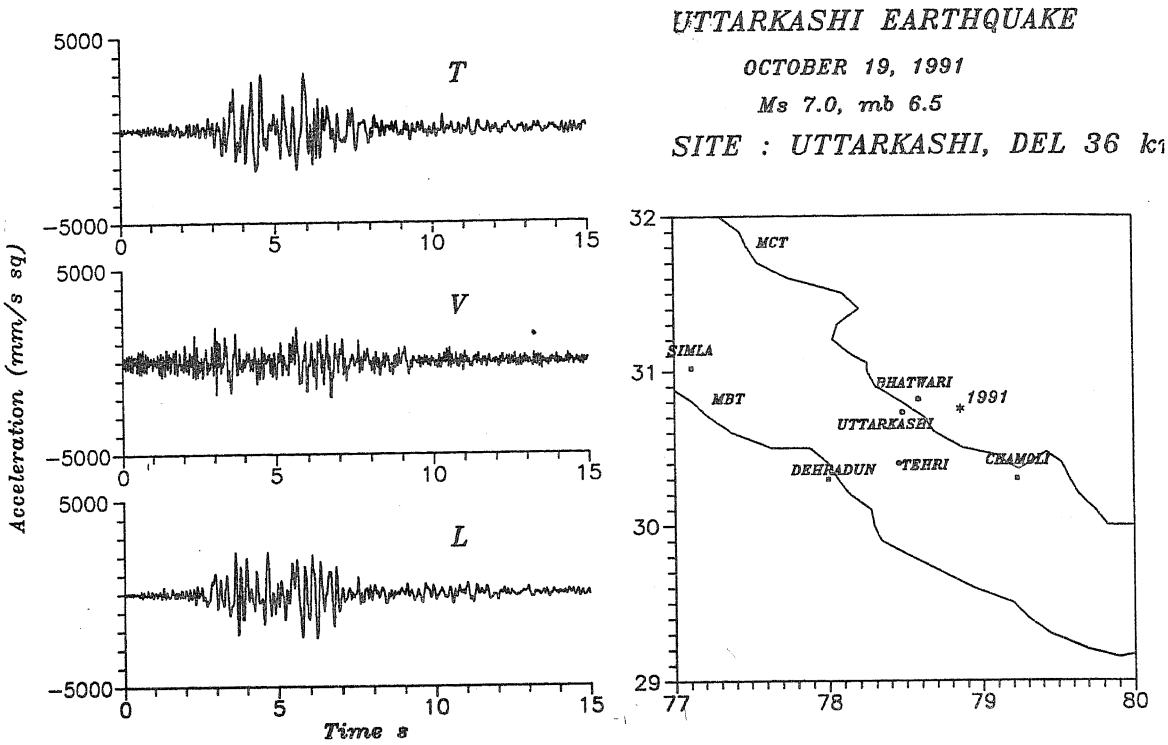


Figure 6. Accelerograms recorded at Uttarkashi for the 1991 Uttarkashi earthquake (M_s = 7).

lar observations for the peak acceleration values. However, the long-period spectral responses begin to show the effects of epicentral distance. For example, at ~ 1 s period, the spectral accelerations at Bhatwari and Uttarkashi are, respectively, 0.22 *g* and 0.10 *g* for the L component, 0.43 *g* and 0.18 *g* for the T component, and 0.11 *g* and 0.10 *g* for the V component. Similarly, at 2 s period, these values are 0.08 *g*, 0.03 *g* for the L component, 0.11 *g* and 0.04 *g* for the T component, and 0.06 *g* and 0.05 *g* for the V component. We notice a falling off of spectral acceleration amplitudes for the horizontal components by a factor of about 3 at these longer periods for a change of epicentral distance from 25 km for Bhatwari to 36 km for Uttarkashi, i.e. an increase of 11 km.

Let us now consider the results obtained for the case of simulation for Tehri. As noted above, peak accelerations of 0.82 *g*, 0.57 *g* and 0.53 *g* have been obtained for T, L and V components, respectively. The largest value is higher by a factor of about two-and-a-half compared to the observed value at Uttarkashi. This increase is theoretically anticipated on account of the smaller distance of Tehri to the fault plane, its smaller epicentral distance (~ 14 km), as well as resonance effects of the near-surface low-velocity layer at Tehri.

The estimated spectral accelerations at Tehri are about 3 to 4 times those observed at Bhatwari for periods up to ~ 1.7 s. At 2 s period, an amplification by 7 has also been observed. In general, the scaling is increasing for longer periods both for the empirical data set (from Uttarkashi to Bhatwari) as well as in the synthetic case (from Bhatwari to Tehri). As discussed above, this is a consequence of increasing coherency of the elastodynamic radiations at longer periods. The scaling by a factor of 3 to 4 is consistent with similar scaling observed between Bhatwari and Uttarkashi for a reduction of 11 km in the epicentral distance. The epicentral distance reduction for Tehri is 10 km with respect to the epicentral distance for Bhatwari. The various data described and discussed above are given in Table 2. The systematic dependence of the spectral accelerations on the epicentral distance discussed above is shown in Figure 8.

The synthetic strong-motion histories are consistent with the observed data at Bhatwari and Uttarkashi for the 1991 Uttarkashi earthquake. They have the appropriate amplitude envelope, duration and frequency characteristics. However, as discussed above, the observations at Bhatwari and Uttarkashi do not provide a suitable basis of estimate of the peak acceleration at Tehri, as both these sites were outside the fault zone; secondly, the Tehri site is well within the fault zone in the model calculations; and thirdly, it includes a low-velocity layer at a shallow depth, which will introduce

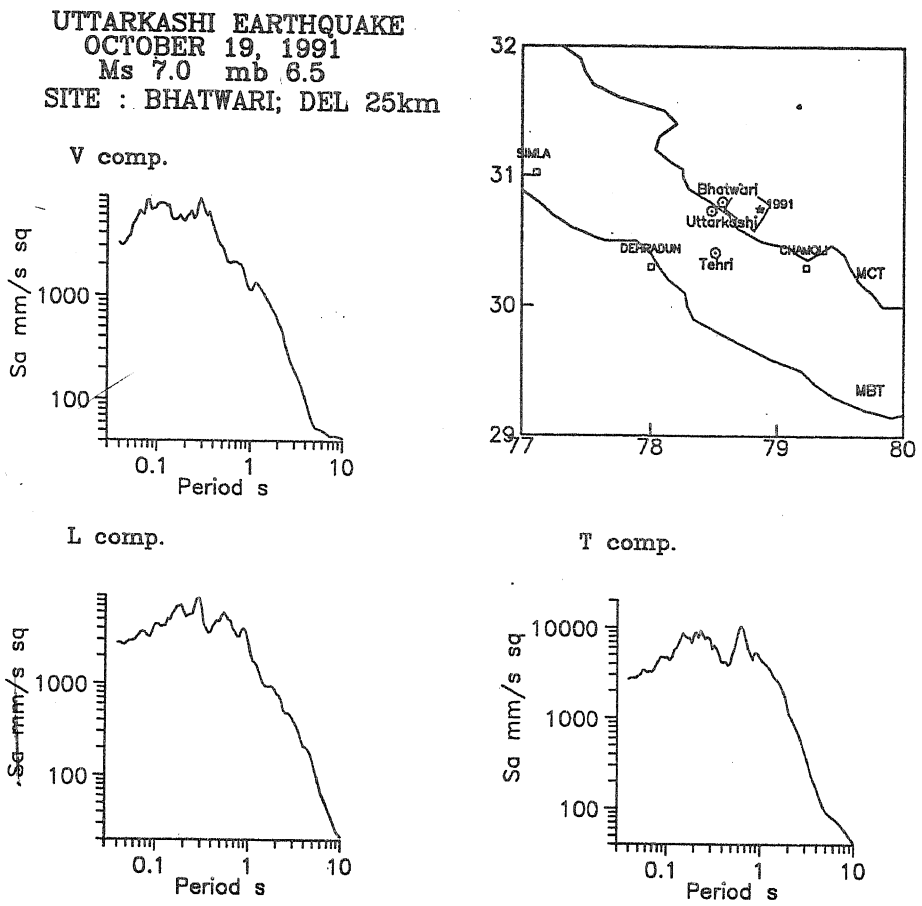


Figure 7. Response spectra at Bhatwari for 5% damping, for the accelerograms in Figure 4.

resonance effects at shorter frequencies. Such a layer is not expected to exist at either Bhatwari or Uttarkashi, both of which are located adjacent to the High Himalaya having a very high rate of erosion. But such a layer is expected to be present at Tehri as indeed brought out from the analysis of the strong motion data for the Uttarkashi earthquake[3]. Furthermore, peak ground accelerations of about 1 g near the faults have been recorded in several earthquakes[8, 9]. Therefore, the maximum peak acceleration of 0.82 g obtained for the T component is a realistic estimate.

The observed data as well as the foregoing analysis indicate that the synthetic accelerograms and the corresponding response spectra for the Tehri site represent a realistic seismic hazard scenario for the type of earthquake and the site under consideration.

Estimation of the most damaging future earthquake in the Tehri area within the next about 100 years

Estimation of seismic hazard in a region depends on the most damaging earthquake that can occur in the region in the time window under consideration. This is illustrated by considering the accelerations expected for large earthquakes. For example, the Idriss[10] formula provides the following estimates for acceleration values for the cases of $M = 8$ and $M = 8.5$ earthquakes at distances of 15 and 25 km (see Table 3).

Table 3 shows two things. First, *there is a drastic increase in the expected acceleration amplitudes as we go from $M = 8$ to $M = 8.5$. Secondly, the high values of the spectral accelerations imply a long duration of shaking with high acceleration values.*

An estimate of the peak acceleration acquires significance when, as is often the case, the seismic response spectrum based on site-specific observational data is not available. In such situations, the amplitudes of the response spectrum are scaled by anchoring the spectral acceleration at zero period of a standard mean level spectra, at the level of the peak acceleration expected for the site, or a parameter based on the same. However, alternative new techniques based on sound theoretical models, such as the one presented here, have now become available to provide estimates of the expected strong ground motion time histories and the corresponding response spectra.

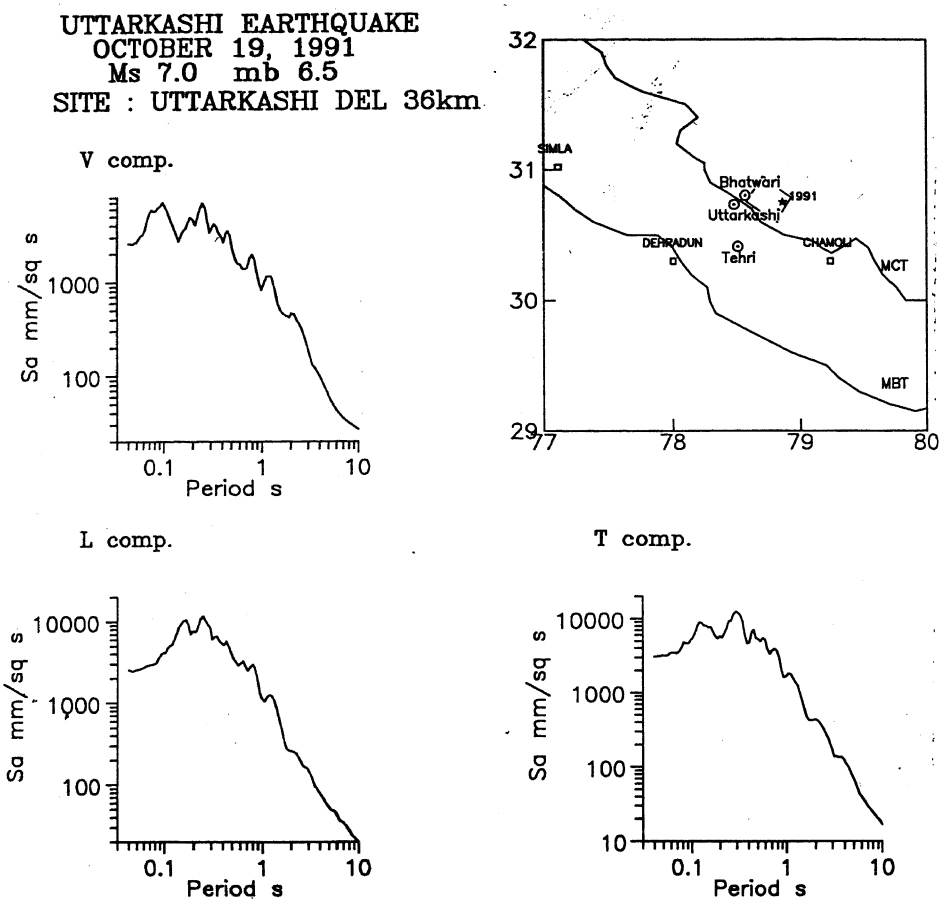


Figure 8. Response spectra at Uttarkashi for 5% damping, for the accelerograms in Figure 5.

The Himalayan belt is one of the most actively deformed zones on earth[12, 13], formed as it is by a succession of processes of strain accumulation and its catastrophic release in the form of earthquakes. The region has experienced four great earthquakes in the past about 100 years, namely, the 1897 Assam earthquake ($M = 8.7+$), the 1905 Kangra earthquake ($M = 8.6$), the 1934 Bihar–Nepal earthquake ($M = 8.4$), and the 1950 Assam earthquake ($M = 8.7$). The occurrence of peak accelerations in excess of $1\,g$ in the great 1897 Assam earthquake has been well documented[14]. Since the geological process of the collision of the Indian plate with the Eurasian plate continues, causing persistent accumulation of strain, it makes certain that such great earthquakes in the Himalaya will continue to occur in the future.

An appraisal of the average return period of the $M = 8+$ earthquake in the region is, therefore, pertinent to the problem of seismic hazard estimation. This can be based on the strain budget of the region. The frequency of great earthquakes is constrained by the rate of strain accumulation in the region. The rate of strain accumulation has been estimated using a variety of techniques. The average rate of plate convergence obtained from marine magnetic anomalies is 5 cm/annum[15]. However, a fraction of this is transferred to the region north of the Himalaya. Thus, it sets an upper limit on the rate of strain accumulation in the region. Using the deformation of the upper Tertiary sediments in the Ganga foredeep, the average rate of strain is estimated to be 1.0–3 cm/annum in the western parts of the Himalaya[16, 17]. In the U.P. part of the Ganga foredeep, the average rate of convergence estimated from the velocity of the southward onlapping of the Indian shield is 1.8 ± 0.4 cm/annum[18]. The elastically stored component of the average rate of convergence estimated by using the rate of seismic moment released by great earthquakes in the Himalaya is 1.7 ± 0.3 cm/annum[19]. This strain becomes available for causing future earthquakes in the region. Using the fractal description of earthquake occurrences together with the complete catalogue of earthquakes, the average rate of slip over the basement thrust in the Himalaya has been estimated to be 3.5 cm/annum[20]. This rate will be approximately equal to the average rate of

Table 2. The values of peak accelerations and S_a (g)

Site	Uttarkashi (observed)			Bhatwari (observed)			Tehri (synthetic)		
Epicentral Distance (km)	36			25			14		
Component	T	V	L	T	V	L	T	V	L
a_p (g)	0.30	0.19	0.24	0.24	0.29	0.25	0.82	0.53	0.57
Factor				~ 1.5	~ 1	~ 1	~ 4	~ 2	~ 2
S_a (g) at T = 0.98 s	0.11	0.16	0.08	0.30	0.47	0.12	0.70	0.96	0.33
Factor				~ 3	~ 3	~ 1.5	~ 2.5	~ 2	~ 3
S_a (g) at T = 1.04 s	0.10	0.18	0.10	0.22	0.43	0.11	0.69	0.98	0.29
Factor				~ 2	~ 2	~ 1	~ 3	~ 2	~ 2.5
S_a (g) at T = 1.23 s	0.12	0.14	0.12	0.16	0.36	0.13	0.61	0.85	0.50
Factor				~ 1	~ 2.5	~ 1	~ 4	~ 2	~ 4
S_a (g) at T = 1.50 s	0.05	0.06	0.05	0.09	0.26	0.09	0.48	0.60	0.41
Factor				~ 2	~ 4	~ 2	~ 5	~ 2	~ 5
S_a (g) at T = 1.70 s	0.03	0.04	0.04	0.09	0.21	0.08	0.49	0.65	0.21
Factor				~ 3	~ 5	~ 2	~ 5	~ 3	~ 3
S_a (g) at T = 2.00 s	0.03	0.04	0.04	0.08	0.11	0.06	0.54	0.76	0.18
Factor				~ 3	~ 2.5	~ 1	~ 7	~ 7	~ 3

'Factor' represents the amplification factor with respect to Uttarkashi in the case of Bhatwari and with respect to Bhatwari in the case of Tehri.

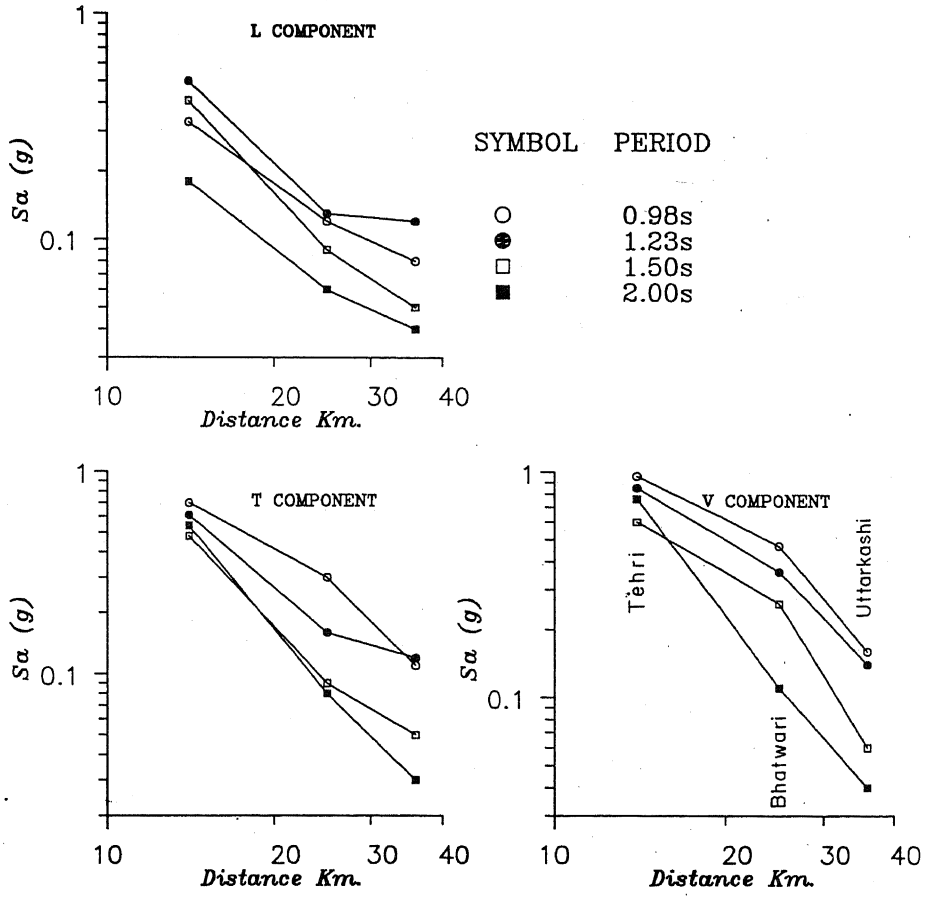


Figure 9. The variation of spectral accelerations as a function of distance and period. The data used are summarized in Table 2.

convergence in the Himalaya. Geodetic measurements using levelling in Dehradun and Kathmandu areas have been analysed to provide an estimate of the average rate of elastic strain accumulation to be 0.7 cm/annum in the outer Himalayan region[21]. The GPS measurements in Nepal have led to an estimate of the convergence rate of 1 ± 1 cm/annum[22].

From the foregoing discussion it is clear that the average rate of convergence in the Himalaya lies in the range of 1–3 cm/annum. The rate may vary along different

Table 3. Acceleration values using Idriss formula[10], quoted from Liam-Finn[11]

	$M = 8$		$M = 8.5$	
	$R = 15$ km	$R = 25$ km	$R = 15$ km	$R = 25$ km
a_p	$0.66\,g$	$0.485\,g$	$0.75\,g$	$0.575\,g$
a_{sp} at $T = 0.3$ s	$2.1\,g$	$1.6\,g$	$2.5\,g$	$2.0\,g$
a_{sp} at $T = 1.0$ s			$1.14\,g$	
a_{sp} at $T = 1.2$ s			$0.96\,g$	
a_{sp} at $T = 1.5$ s			$0.75\,g$	
a_{sp} at $T = 2.0$ s	$0.44\,g$	$0.31\,g$	$0.57\,g$	$0.41\,g$

sections of the Himalaya. This can be revealed by detailed investigations using modern techniques like the GPS.

If one considers the average rate of elastically stored convergence to be 2 cm/annum then sufficient strain will become available in any sector of the Himalaya to generate a great earthquake in a period of about 500 years. In the sector in front of the Salt range, where the rate has been estimated to be 1 cm/annum, the average return period will be correspondingly longer. In the sector containing the Pir Panjal Range, where the rate of convergence is estimated to be 3 cm/annum, the average return period is anticipated to be shorter. In the central gap region the estimated rate of convergence is close to the average value. Therefore, the 500-year return period is quite representative.

The four great earthquakes in the Himalaya over the past 100 years have between them relieved the accumulated strain in about 35% of the Himalayan plate boundary. The remaining sections of the Himalayan plate boundary have not experienced a great earthquake for the past at least 300 years, possibly 500 years or more. Thus, these sections, constitute veritable seismic gaps, with a large amount of accumulated strain, and are, therefore, places of high seismic potential for a future great earthquake in the next 100–300 years or so. The Tehri dam site lies in one such gap (Central gap)[23, 24], for which the shorter time interval applies.

It is also relevant to consider peak accelerations in the Himalaya, estimated using the standard probabilistic seismic hazard analysis[25]. The peak accelerations expected in the Lesser Himalaya (where the dam is sited) are $0.7^{+}g$ for a one in ten chance of exceedance in any time window of 50 years. When this is translated to a 100-year exposure period, the a_p value scales to about $0.9\,g$ (ref. 26).

The analysis of both the seismological and the strong-motion data are consistent in estimating that for the case of a great earthquake both the peak and the spectral accelerations ($T \sim 1$ s) in the fault zone in the Tehri region are expected to be around $0.9\,g$.

Conclusions

Recapitulating, the following conclusions emerge from the present investigation.

1. The observational strong-motion data for the 1991 Uttarkashi earthquake ($M_s = 7$) show that spectral accelerations of $0.43\,g$ ($T = 1.04$ s), $0.36\,g$ ($T = 1.23$ s) and $0.26\,g$ ($T = 1.5$ s) have been recorded at an epicentral distance of 25 km.

2. The corresponding synthetic strong motions for a hypothetical $M_s = 7$ earthquake at an epicentral distance of 14 km and located within the surface projection of the fault zone are $0.98\,g$ ($T = 1.04$ s), $0.85\,g$ ($T = 1.23$ s) and $0.60\,g$ ($T = 1.5$ s).

3. The spectral accelerations in the 1–2 s period range, estimated from model simulation at Tehri for the case of an $M_w = 8.5$ earthquake situated under Tehri at a depth of 12 km, are $\sim 1\,g$ (refs. 7, 8).

4. The Idriss[10] as well as the Campbell[27] empirical formulae also give spectral accelerations ($T = 2$ s) at 15 km distance to be at least $\sim 0.57\,g$ for an $M = 8.5$ earthquake. The acceleration will be higher at a closer distance of 12 km.

5. The rate of strain accumulation in the Central Gap (where Garhwal is situated) is around 2 cm/annum, which provides enough strain accumulation to generate a great earthquake in about 500 years. Since there has been no great earthquake in this gap for at least about 300 years and perhaps 500 years, the region (Central Gap) must be ripe for a great earthquake in the next 100 years or so. Of course, it is possible that the accumulated strain is initially released partially in a number of $M = 7$ or so earthquakes, instead of a single great earthquake ($M > 8$). In this case the occurrence of the great gap-filling earthquake will at best be somewhat delayed, but not forestalled. However, the scenario historically repeated by all great earthquakes in the Himalaya is the most appropriate basis for estimating seismic hazard.

6. The probabilistic seismic hazard analysis of the Himalaya shows that in any 100-year time window the probability of exceeding a peak acceleration of $\sim 0.9\,g$ in the Garhwal Himalaya is 0.1.

7. The above estimates of the seismic hazard characterization are consistent with the currently available corpus of geological, geophysical, seismological and earthquake strong-motion data. Therefore, it would be prudent to heed these estimates while scientifically evaluating the soundness of the aseismic design of large critical structures in the region.

8. There is a very high level of seismic hazard in the seismic gaps and in adjacent areas to their south with their densely populated centres. A systematic implementation of a well laid-out hazard mitigation plan is urgently called for.

1. Zeng, Y., Anderson, J. G. and Yu, G., *Geophys. Res. Lett.*, 1993, **21**, 725–728.
2. Khattri, K. N., Yu, G., Anderson, J., Brune, J. and Zeng, Y., *Curr. Sci.*, 1994, **67**, 343–353.
3. Yu, G., Khattri, K. N., Anderson, J., Brune, J. and Zeng, Y., *Bull. Seismol. Soc. Am.*, 1995, **85**, 31–50.
4. Brune, J. N., *J. Geophys. Res.*, 1970, **75**, 4997–5009.
5. Brune, J. N., *J. Geophys. Res.*, 1971, **76**, 5002.
6. Frankel, A., *J. Geophys. Res.*, 1991, **96**, 6291–6302.
7. Kasahara, K., *Earthquake Mechanics*, Cambridge University Press, New York, 1981, pp. 19, 56–57.
8. Brune, J. N., *Tectonophysics*, 1993, **218**, 281–286.
9. Ambraseys, N. N., in *Earthquake Hazard and Large Dams in the Himalaya* (ed. Gaur, V. K.), INTACH, New Delhi, pp. 111–115.
10. Idriss, I. M., Selection of earthquake ground motions at rock sites, Report prepared for the Structures Division, Building and Fire Research Laboratory, National Institute of Standards for Technology, Department of Civil Engineering, University of California, Davis, 1991.
11. Liam-Finn, W. D., *Earthquake Hazard and Large Dams in the Himalaya* (ed. Gaur, V. K.), INTACH, New Delhi, pp. 116–136.
12. Valdiya, K. S., in *Earthquake Hazard and Large Dams in the Himalaya* (ed. Gaur, V. K.), INTACH, New Delhi, pp. 1–34.
13. Nakata, T., in *Tectonics of the Western Himalayas* (eds Malinconico, L. L. (Jr) and Lillie, R. J.), Geol. Soc. America Spl. Paper, 232, pp. 243–264.
14. Oldham, R. D., *Mem. Geol. Surv. India*, 1899, **26**, 1–379.
15. DeMets, C., Gordon, R. G., Argus, D. F. and Stein, S., *Geophys. J. Int.*, 1990, **101**, 425–478.
16. Burbank, D. W. and Raynolds, R. H. G., *Nature*, 1984, **311**, 114–118.
17. Baker, D. M., Lillie, R. J., Yeats, R. S., Johnson, G. D., Yousuf, M. and Zamin, A. S. H., *Geology*, 1988, **16**, 3–7.
18. Lyon-Caen, H. and Molnar, P., *Tectonics*, 1985, **4**, 513–538.
19. Molnar, P. and Pandey, M. R., *Proc. Indian Acad. Sci. (Earth and Planet. Sci.)*, 1989, **98**, 25.
20. Khattri, K. N., *Curr. Sci.*, 1995, in press.
21. Chander, R. and Gahalaut, V. K., *Curr. Sci.*, 1994, **67**, 531–534.
22. Bilham, R. and Jackson, M., *J. Nepal Geol. Soc.*, 1994, **10**, 18.
23. Khattri, K. N. and Tyagi, A. K., *Tectonophysics*, 1983, **96**, 281–297.
24. Khattri, K. N., *Tectonophysics*, 1987, **138**, 79–92.
25. Khattri, K. N., Rogers, A. M., Perkins, D. M. and Algermissen, S. T., *Tectonophysics*, 1984, **198**, 93.
26. Khattri, K. N., Rogers, A. M. and Perkins, D. M., *Bull. ISET*, 1983, **20**, 1–22.
27. Campbell, K. W., *Himalaya with References to Tehri Dam, in Earthquake Hazard and Large Dams in the Himalaya* (ed. Gaur, V. K.), INTACH, New Delhi, pp. 93–103.

ACKNOWLEDGEMENTS. This research was supported by the Emeritus Scientist Scheme of the CSIR and the WIHG, who have kindly hosted me. I am indebted to Y. Zeng, G. Yu, and J. Anderson for allowing the use of their codes. Thanks are due to G. Yu, who kindly ran some simulations, and to Sri Ram for running some graphics. I have benefited immensely from discussions with Professors V. K. Gaur, R. Chander and R. N. Iyengar.

Received 12 May 1995; revised accepted 17 May 1995

RESEARCH COMMUNICATIONS

Sr isotopes in rivers of India and Pakistan: A reconnaissance study

J. R. Trivedi, Kanchan Pande, S. Krishnaswami and M. M. Sarin

Physical Research Laboratory, Ahmedabad 380 009, India

It is well established[1] that the isotopic composition of Sr in sea water has oscillated on different time scales during the past ~600 Ma. These oscillations result from changes in the relative contributions of dissolved Sr to sea water from its two major sources, continental weathering and hydrothermal interactions, with their characteristic $^{87}Sr/^{86}Sr$. The continental input via fluvial weathering is generally more radiogenic (average $^{87}Sr/^{86}Sr$ 0.7119 ref. 2) compared to the hydrothermal $^{87}Sr/^{86}Sr$ of 0.7035 (ref. 2). The Sr flux and $^{87}Sr/^{86}Sr$ of different river systems vary widely[2–4] and in some way are related to the lithology of the drainage basin, the intensity of weathering and the efficiency with which they transport their dissolved load. We have measured the dissolved Sr concentrations and $^{87}Sr/^{86}Sr$ of the major rivers draining India and Pakistan. These measurements provide data on the flux of Sr isotopes transported by these rivers to the ocean and their impact on the Sr isotope evolution of the oceans. Further, as the river basins sampled comprise a diverse set of lithologic, tectonic and climatic regimes, they provide an opportunity to study the controls of these factors place in determining the Sr isotopic systematics of rivers.

THIS work forms a part of our geochemical reconnaissance study of the Indian rivers. Major ions, Sr and U isotopes were the principal constituents measured in the samples, the results of some of our earlier work are published elsewhere[5–9]. Figure 1 shows the sampling sites.

The rivers sampled can be conveniently classified into two groups[10,11] in terms of their basin geology, tectonics and climate – the peninsular and the extrapeninsular (Himalayan). The peninsular region is a tectonically quiescent shield area that has several mountain systems such as the Aravallis, the western and eastern Ghats and the Nilgiris, most of which are geologically mature. This shield is made of Pre-Cambrian rocks of diverse origin,

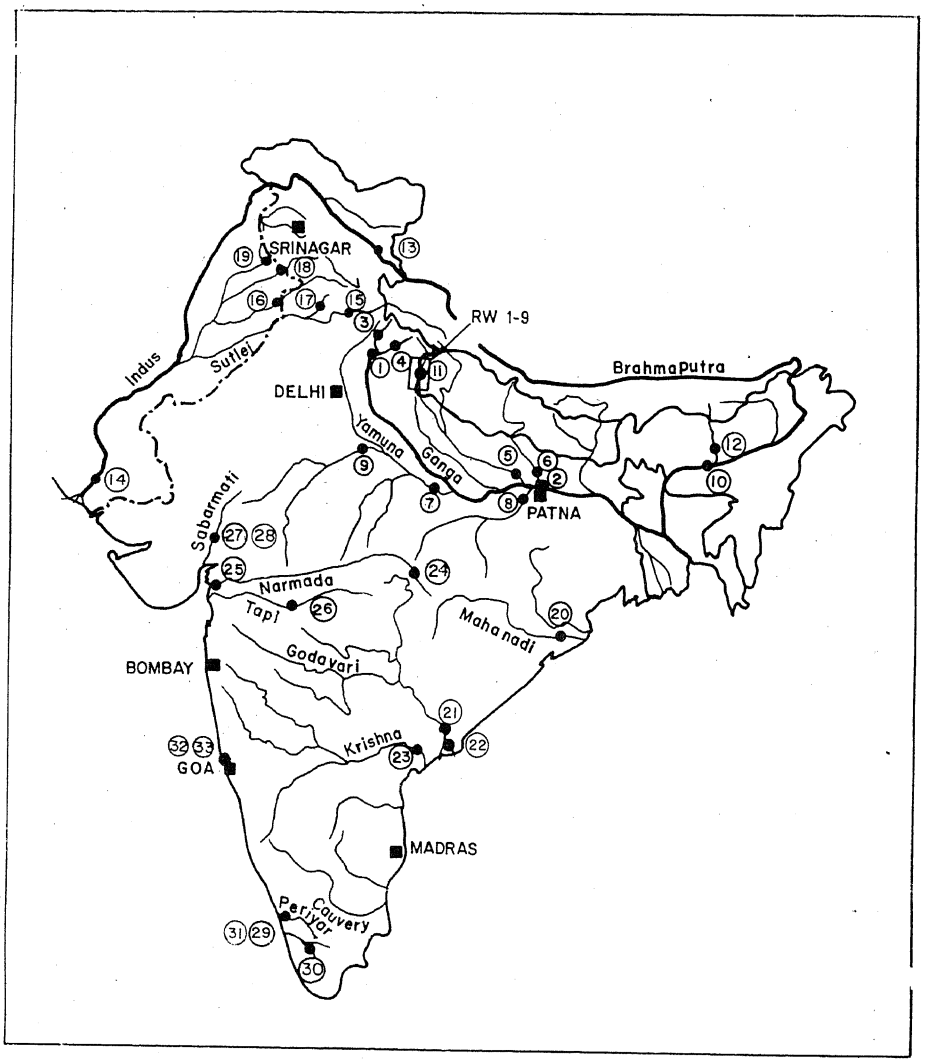


Figure 1. Sampling locations of rivers from India. The sampling location of Indus and its tributaries are given in ref. 9. The locations of samples from the Kali–Sarju system (RW 1–9) are marked by rectangle.

most of which have been greatly modified and metamorphosed. These ancient rocks (gneisses, granite, charnokites, greenstones, granulites, etc.) are overlain in many areas by Pre-Cambrian and later sedimentaries (Cuddapah, Vindhyan, Delhis and Gondwanas) and by extensive sheets of lavas of the Cretaceous/Tertiary Deccan traps. Some Mesozoic and Tertiary sediments are also found along the coastal regions[10,11].

There are a number of rivers traversing the peninsular India, the larger ones being Mahanadi, Godavari, Krishna, Pennar and Cauvery, which flow eastward into the Bay of Bengal, and Narmada, Tapi, Sabarmati and Periyar, flowing westward into the Arabian sea. Banas, Luni, Chambal, Sindh, Betwa and Son are the peninsular rivers (lowland rivers[5]) belonging to the Ganga system.

Similar to the peninsular region, the extrapeninsular (Himalayan) region also contains some Pre-Cambrian rocks. These old rocks in the Lesser and Greater Himalayas are comprised predominantly of granitic gneisses and crystallines[12,13]. In the Lesser Himalaya there are also extensive exposures of sedimentary sequences (carbonates, sandstones, shales). The peninsular and extrapeninsular regions are separated by the Indo-Gangetic alluvial plain consisting of massive beds of clay, sand and gravel. The clay zones of alluvial plains contain impure calcareous material in the form of irregular concretions called 'Kankar'. The calcareous material at places can be as abundant as 30%. A large part of the soils is impregnated by alkaline and saline salts[5]. In contrast to the peninsular shield area, the Himalayas are

Table 1. K, Ca, Rb and Sr isotopes in rivers from India–Pakistan

#	River	Location	K (μmol/l)	Ca (μmol/l)	Rb (nmol/l)	Sr (nmol/l)	$^{87}Sr/^{86}Sr$
Himalayan rivers							
Ganga-Brahmaputra system							
1.	Ganga	Rishikesh	38	353	29	580	0.7365
2.	Ganga	Patna	63	547	11	1017	0.7249
3.	Bhagirathi	Devprayag	51	385	54	529	0.7589
4.	Alaknanda	Bagwan	51	429	40	644	0.7385
5.	Ghaghara	Ayodhya	59	528	23	756	0.7349
6.	Gandak	Hajipur	84	649	58	847	0.7400
7.	Yamuna	Allahabad	57	547	7	1509	0.7149
8.	Son	Ara	40	434	6	644	0.7200
9.	Chambal	Dholpur	53	829	5	2290	0.7127
10.	Brahmaputra	Goalpara	46	318	32	670	0.7197
11.	Kali	Jauljibi	39	551	70	1320	0.7276
Indus system							
13.	Indua	Loma	34	522	41	1394	0.7104
14.	Indus	Thatta	109	957	10	3689	0.7111
15.	Sutlej	Salapar	82	622	79	1962	0.7166
16.	Ravi	Lahore	71	1070	30	1858	0.7295
17.	Beas	Hanogi Mata	25	148	31	169	0.7448
18.	Chenab	Gujrawala	63	1050	32	2155	0.7218
19.	Jhelum	Lala Musa	40	950	13	2839	0.7125
Peninsular rivers							
20.	Mahanadi	Cuttak	40	303	11	704	0.7193
21.	Godavari	Dhavleswaram	59	350	24	1179	0.7163
22.	Godavari	Ravulapalam	63	400	26	1375	0.7152
23.	Krishna	Vijaywada	77	600	20	3742	0.7142
24.	Narmada	Amarkuntak	18	190	13	316	0.7144
25.	Narmada	Upstream Bharuch	46	742	6	2156	0.7114
26.	Tapi	Near Bhusaval	–	585	14	2083	0.7097
27.	Sabarmati	Kheroj	120	908	15	3358	0.7177
28.	Sabarmati	Sijaria	77	1388	10	5199	0.7159
29.	Periyar	Alwaye	(920)	90	36	365	0.7194
30.	Pambayar	Tallawadi	28	43	35	180	0.7230
31.	Achankovil	Pathanamthitta	26	63	26	290	0.7149
32.	Mandovi	Near Valpoi	26	242	24	352	0.7196
33.	Zuari	Quepem	36	148	32	431	0.7235
8.	Son	Ara	40	434	6	644	0.7200
9.	Chambal	Dholpur	53	829	5	2290	0.7127

Data for Ganga–Brahmaputra from ref. 7 and for the Indus system from ref. 9 and for the peninsular rivers this study.
#Denotes sample location in Figure 1.

Table 2. K, Ca, Rb and Sr isotopes in the Kali–Sarju system (headwaters of Ghaghra)

# Code	River	Location	K (μmol/l)	Ca (μmol/l)	Rb (nmol/l)	Sr (nmol/l)	$^{87}Sr/^{86}Sr$
RW-1	Sarju	Sheraghat	33	577	37	263	0.7585
RW-2	Ramganga	Thal	30	379	41	296	0.7757
RW-3	Ghat Gad	Near Patet	34	609	35	316	0.7508
RW-4	Kali	Jauljibi	39	551	70	1320	0.7276
RW-4 (R*)					–		0.7277
RW-5	Gori	Baram	41	456	63	768	0.7490
RW-6	Sarju	Rameshwar	34	551	44	311	0.7514
RW-7	Sarju	Near Rameshwar	33	551	36	286	0.7553
RW-8	Panar	Near Rameshwar	52	545	22	1245	0.7408
RW-9	Suyal	Lohali	39	211	17	413	0.7351

*Repeat analysis of RW-4.

tectonically active, still rising and being extensively eroded by the fast-flowing rivers. Glaciers abound the higher reaches and also contribute to the accelerated weathering. The lithology of the suture zone in the Trans-Himalaya is dominated by granites and sediments[9,11].

Ganga, Brahmaputra and Indus are the major river systems of the extrapeninsular region, the first two flowing eastward into the Bay of Bengal and the last into the Arabian sea. Ganga and its tributaries from the Himalaya, Yamuna, Kali, Gandak, Kosi and Ghaghara, drain the Pre-Cambrian crystallines and sedimentary sequences in the southern slopes of the Himalaya, whereas Brahmaputra flows along the Indus Tsangpo suture zone in Tibet before taking a turn at the syntaxial bend and entering the crystalline terrains of the Himalaya[5]. The main tributaries of the Indus system – Sutlej, Beas, Ravi, Chenab (Chandra and Bhaga) and Jhelum – drain the crystalline terrains of the Himalaya in their upper reaches (similar to Ganga and its tributaries), whereas Indus and its northern tributaries, Nubra and Shyok, drain a suture zone lithology in a monsoon shadow region[9].

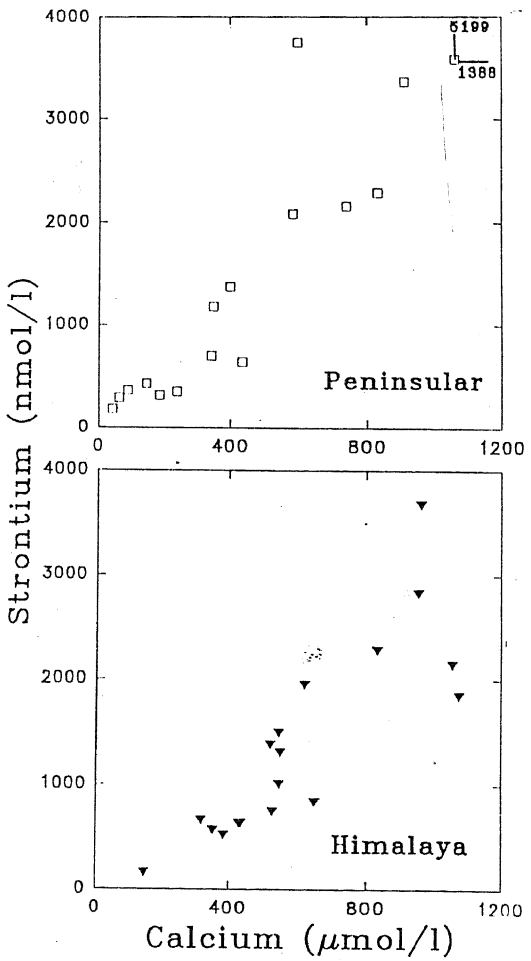


Figure 2. Variation of Sr concentration with calcium in the peninsular and Himalayan rivers. The Sr concentrations generally increase with Ca.

The peninsular rivers are seasonal (monsoonal) in contrast to the Himalayan ones, which are fed by glaciers and are perennial, though their water flux increases considerably during monsoon.

Most of the rivers were sampled during the past several years. The sampling of the Kali–Sarju system (headwaters of the Ghaghra) was carried out recently. The sampling procedures are discussed in detail elsewhere[5]. Briefly, the river waters were filtered through 0.4 μm Nuclepore filters within 5–6 h of collection, acidified to pH~2 with HNO_3 and stored in polyethylene bottles till the time of measurement. Rb and Sr isotope measurements were made in about 50 g of water following isotope dilution mass spectrometric procedures[14]. The weighed water samples were dried with HF/HNO_3, put in HCl and spiked with ^{87}Rb and ^{84}Sr tracers, and dried. The dried residue was taken in 2.5 N HCl and the Rb and Sr fractions were separated on a cation exchange resin (Dowex 50×8, 200–400 mesh). The Rb and Sr fractions were separately dried, taken in a drop of orthophosphoric acid and evaporated directly onto an outgassed tantalum filament of the ion source. The samples were run on a 9″ radius single-focusing mass spectrometer fitted with a Faraday cup collector and the ion current was recorded digitally.

The Sr isotopic ratios were normalized to $^{86}Sr/^{88}Sr = 0.1194$. The long-term stability of the mass spectrometer was checked by measuring the ratio $^{87}Sr/^{86}Sr$ of the NBS-987 standard periodically. The mean $^{87}Sr/^{86}Sr$ of the standard based on nine runs during 1991–93 is 0.71025 ± 0.00005. In addition, to check on the accuracy of $^{87}Sr/^{86}Sr$ measurements we analysed a sea water sample from the Arabian sea and $CaCO_3$ from a living coral from the Lakshadweep. For these two samples $^{87}Sr/^{86}Sr$ are 0.7097 ± 0.0006 and 0.7092 ± 0.0002, respectively, in good agreement with contemporary sea water values. The total Rb and Sr blanks from reagents and chemical procedures are ~5 and ~50 pmol, respectively, much lower than the typical values of about 1 and 25 nmol for Rb and Sr in 50 g water samples. The overall precision on $^{87}Sr/^{86}Sr$ (two standard errors of the mean of about 50 ratio measurements) is about ± 0.0005. These errors are considerably smaller than the scale of Sr isotope variations observed in the river waters analysed. Ca and K measurements were made by atomic absorption spectrophotometry[5].

The dissolved Rb and Sr concentrations and $^{87}Sr/^{86}Sr$ of the water samples analysed are given in Table 1 along with their K and Ca abundances. Table 2 lists the same for the Kali-Sarju system (also known as Mahakali–Sarju) headwaters of Ghaghara, a major tributary to the Ganga. More detailed data on Sr isotopes, Rb and major

ions for the rivers belonging to the Ganga–Brahmaputra and Indus systems are published elsewhere[7,9].

The Kali–Sarju waters were sampled to test if $^{87}Sr/^{86}Sr$ for the headwaters of major tributaries of Ganga, are as radiogenic as of Bhagirathi and the Alaknanda, sampled by us earlier[7]. The results show that the Sr isotope ratios in all these waters are highly radiogenic, and as will be discussed later, such high values probably result from weathering of rich in radiogenic Sr silicates in the drainage basin.

A cursory look at the data in Table 1 reveals that the Sr concentration in the samples ranges between 169 and 5199 nmol/l, similar to that observed for rivers in other parts of the world[2-4].

The Sr concentration shows a general increase with Ca (Figure 2), the correlation being relatively tight for rivers from the peninsular region. The Sr/Ca molar ratios (nmol Sr/μmol Ca) of the peninsular and the Himalayan rivers overlap within the scatter of the data. The mean molar Sr/Ca ratio corresponds to a Sr/Ca weight ratio (g/g) of $\sim 7 \times 10^{-3}$ and $\sim 5 \times 10^{-3}$ for the peninsular and the Himalayan rivers, respectively. The Sr/Ca ratio in rivers is determined by the supply of Sr and Ca to the waters by various rocks and minerals, and their mixing proportions. An important source of Ca to these waters (Table 1) is the weathering of carbonates[5–9]. The Sr/Ca weight ratio in typical carbonates is $1-2 \times 10^{-3}$ (ref. 15). Congruent weathering of carbonates (for Ca and Sr) would yield Sr/Ca ratio in river waters significantly lower than that we have measured. Such an observation (Figure 2) would require that either Sr is mobilized preferentially over Ca from carbonates such that the Sr/Ca ratio in rivers is higher than that in carbonates or that these rivers receive Sr from other sources such as evaporites and silicates from the upper crust, which have Sr/Ca higher than that in typical carbonates. Basalts, granites and shales have Sr/Ca about 3 to 10 times higher than that in carbonates. As discussed below, $^{87}Sr/^{86}Sr$ for these river waters also requires that they derive some of the Sr from a more radiogenic source than typical carbonates, such as silicates of the upper crust.

$^{87}Sr/^{86}Sr$ for the rivers sampled (Table 1, Figure 3) lies in the range 0.7097–0.7589; almost all samples have $^{87}Sr/^{86}Sr$ exceeding the global average estimate of 0.7119 (ref. 2). The data also show that $^{87}Sr/^{86}Sr$ for the peninsular rivers falls within a narrow range of 0.71–0.72 and is generally lower than that for the Himalayan rivers draining the granite/gneiss crystallines. The contrast is more pronounced when the comparison is made with that for the headwaters of Ganga[7] and Ghaghara (Table 2, Figure 3). All the headwater samples analysed, except Kali, have $^{87}Sr/^{86}Sr$ exceeding 0.73, with Sr concentration in the range of 140–1390 nmol/l (Table 2, ref. 7).

$^{87}Sr/^{86}Sr$ of dissolved Sr in rivers is governed by the magnitude of Sr contribution from various source rocks

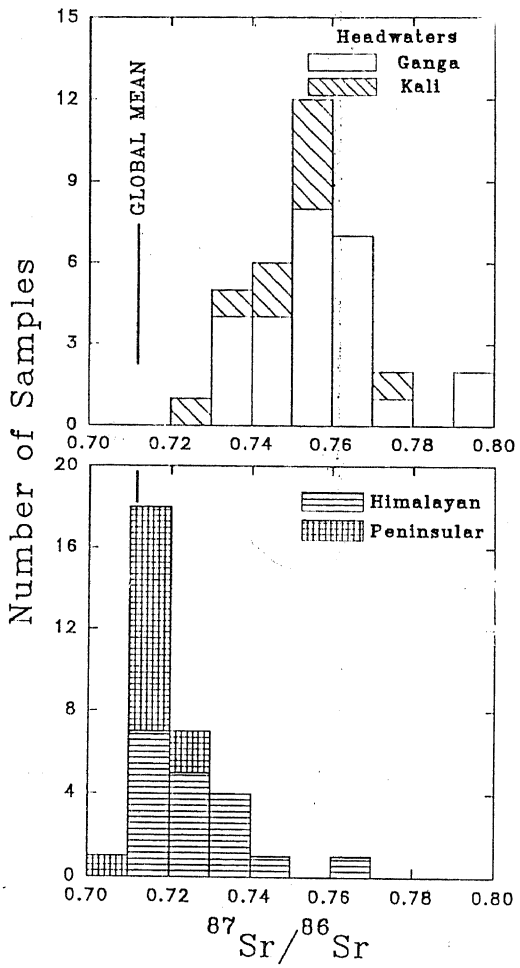


Figure 3. Frequency distribution of dissolved $^{87}Sr/^{86}Sr$ in the rivers sampled. Data for Ganga headwaters from ref. 7 and Kali River (this study). For comparison the data for Himalayan and peninsular rivers from Table 1 are also shown. The distinctly higher radiogenic $^{87}Sr/^{86}Sr$ of the headwaters is evident.

(e.g. carbonates, silicates, evaporites) to the effluent waters and their respective $^{87}Sr/^{86}Sr$. For typical carbonates and evaporites this ratio is constrained within a narrow range of 0.706 – 0.709, whereas that for silicates it can vary widely both within and among various drainage basins. Basaltic terrains set a lower limit on $^{87}Sr/^{86}Sr$ for crustal silicates with values of ~ 0.705. The upper limit on $^{87}Sr/^{86}Sr$ in silicates is less constrained, though values exceeding 1.0 are encountered in individual rock suites, such high values are unlikely to be representative of river basins as a whole.

For all the river waters sampled in this study (Tables 1 and 2) $^{87}Sr/^{86}Sr$ is higher than that for typical marine carbonates, suggesting that at least a part of the Sr in

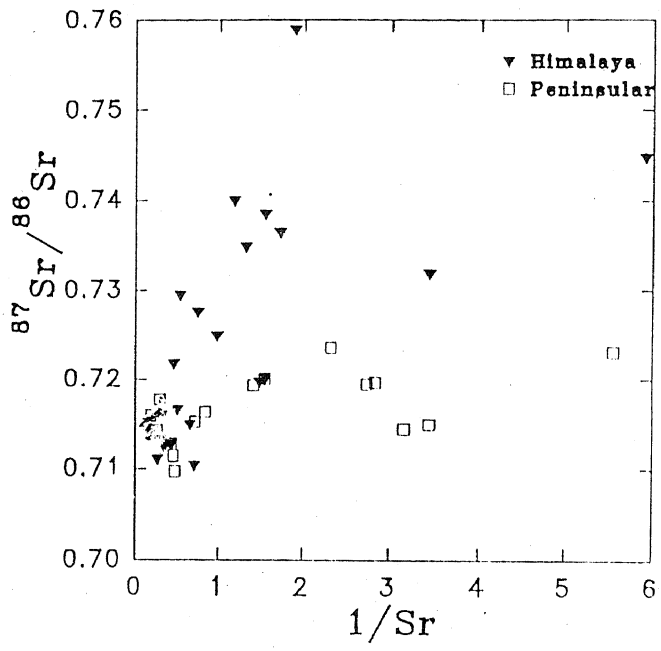


Figure 4. Sr isotope systematics in the rivers sampled. The $^{87}Sr/^{86}Sr$ increases sharply with decreasing Sr concentration in the Himalayan rivers. The peninsular rivers also show a similar trend, but with a gentler slope.

these rivers is derived from a source having more radiogenic $^{87}Sr/^{86}Sr$. The plot of $^{87}Sr/^{86}Sr$ versus 1/Sr (Figure 4) reveals that for the Himalayan rivers there is a steep increase in $^{87}Sr/^{86}Sr$ as the Sr concentration decreases. The peninsular rivers also seem to show an increase in $^{87}Sr/^{86}Sr$ as Sr decreases, but with a much gentler slope. The broad trend in Figure 4 can be interpreted in terms of mixing of Sr between a high $^{87}Sr/^{86}Sr$ end member with a low $^{87}Sr/^{86}Sr$ component (e.g. typical carbonates/evaporites). The relation between $^{87}Sr/^{86}Sr$ and Sr concentration for the Himalayan rivers, particularly those belonging to the Ganga system, are distinctly different from other global rivers, in accordance with our earlier observations and other studies[2,7,9,16,17]. Many of these rivers are characterized by high $^{87}Sr/^{86}Sr$ and high Sr concentrations (Figure 4). The source of such high $^{87}Sr/^{86}Sr$ in these rivers is not clear. Weathering of Pre-Cambrian granites and gneisses[7,9] and rocks and minerals enriched in radiogenic Sr due to regional metamorphism, such as the leucogranites and gneisses[16] and carbonates[17], have been suggested as possible sources of high $^{87}Sr/^{86}Sr$ in these rivers. Studies[18] of $^{87}Sr/^{86}Sr$ in sediments from the Bay of Bengal indicate that the dominant source of these sediments for the past ~17 Ma is the higher Himalayan crystallines. Available data from the drainage basin of the Ganga headwaters and its tributaries show that whole-rock $^{87}Sr/^{86}Sr$ are quite radiogenic, with the bulk of the samples clustering around 0.76 and a significant fraction of the samples having values in excess of 0.90 (Figure 5). Such high $^{87}Sr/^{86}Sr$ in the silicates can account for the high $^{87}Sr/^{86}Sr$ in the headwaters of the Ganga system. However, data on Sr isotope composition of carbonates from the drainage basin are needed to confirm this. Whatever be the source of $^{87}Sr/^{86}Sr$ to Ganga, its high value has considerably influenced the Sr isotope evolution of the oceans since the Cenozoic[7].

The data in Figure 4 show an overall increasing trend in $^{87}Sr/^{86}Sr$ with decreasing Sr, but there is a significant scatter. The scatter within each trend probably reflects heterogenity in the Sr isotope composition of the silicate end member. Such a heterogenity can arise as the Himalayan and the peninsular regions are large geographical areas having lithologically distinct sub-basins which the different rivers drain. The wide range of $^{87}Sr/^{86}Sr$ for the silicates from the drainage basin of the Ganga system headwaters (Figure 5) attests to this hypothesis.

The Sr isotope composition of the peninsular rivers is similar to that reported for rivers from other continental shield areas. The source of high radiogenic Sr to Zuari and Mandovi rivers could be the granites and gneisses in their basin, some of which have $^{87}Sr/^{86}Sr$ of 0.713–1.13 (ref. 19). Periyar and Pambayar flow mainly through granulites; the Sr isotopic analysis on these rocks are meagre, and those which are available are in the range 0.78–0.89 (refs. 20 and 21). Mahanadi and Godavari traverse through Deccan trap basalts, unmetamorphosed sedimentary sequences and the high-grade metamorphic terrain of the eastern Ghats. The measured $^{87}Sr/^{86}Sr$ for these river samples integrates the contributions from all these lithologies. Tapi, flowing mainly through the Deccan trap basalts, has the lowest $^{87}Sr/^{86}Sr$ among all the samples (Table 1).

Based on the Sr concentration and $^{87}Sr/^{86}Sr$ in the peninsular rivers (Table 1) and those reported earlier by us[7,9] for Ganga, Brahmaputra and Indus, the discharge-weighted mean Sr concentration in these rivers is calculated to be ~ 1.5 μmol/l, with a $^{87}Sr/^{86}Sr$ of 0.716. The mean Sr concentration in the rivers of India–Pakistan is higher than the global average value of 0.89 μmol/l[-2]; our sampling of some of the rivers (e.g. Indus) during their lean flow may be a likely cause for the high mean concentration. The discharge-weighted Sr/Ca molar ratio of these rivers is 2.7×10^{-3}, which compares with a value of $\sim 2.2 \times 10^{3}$ for the world rivers (based on Sr and Ca concentrations of 0.89 and 398 μmol/l, respectively). The annual flux of Sr transported by the rivers from India–Pakistan to the ocean is ~2200 million moles if our measurements can be considered representative of these rivers. These calculations show that the rivers from India–Pakistan account for ~6.5% of the global riverine Sr flux to the oceans, with a $^{87}Sr/^{86}Sr$ marginally more radiogenic than the global average.

The influence of anthropogenic effects in modifying the natural Sr flux and its isotopic composition is diffi-

Table 3. Rb, Sr concentrations and $^{87}Sr/^{86}Sr$ in fertilizers

Code	Type	Source	Rb (ppm)	Sr (ppm)	$^{87}Sr/^{86}Sr$*
NPK	Potassium ammonium phosphate	RAMBAN Co. (Dehradun)	24	15	0.7087
MOP	Muriate of potash	–	122	6.4	0.7494
DAP	Ammonium phosphate	SPIC (Dehradun)	2.0	130	0.7080
GUUREA	Urea	KRIBHCO (Ahmedabad)	3.1	0.05	–
NBPHOS	Ammonium phosphate	NARMADA Fertilizers (Ahmedabad)	1.3	554	0.7080
GUJDAP	Ammonium phosphate	IFFCO	3.6	12	0.7091

*Error ($\pm 2\sigma$) is ~0.0005

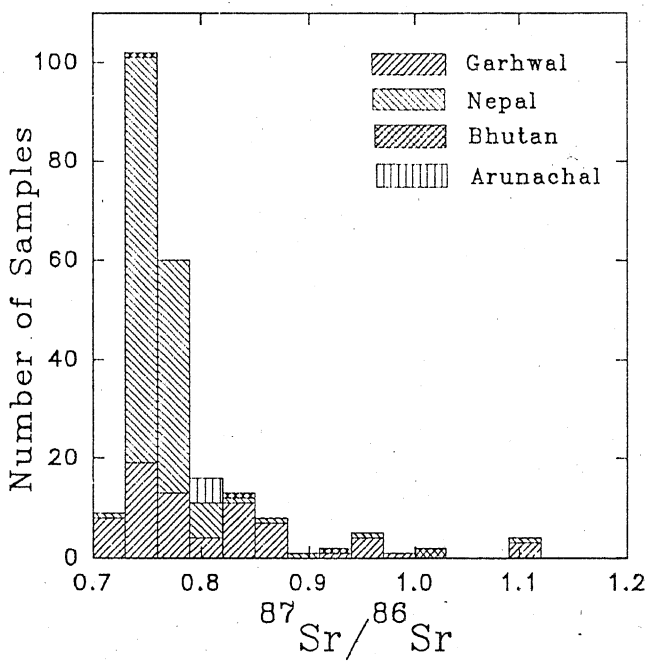


Figure 5. Frequency distribution of $^{87}Sr/^{86}Sr$ in whole-rock samples from the drainage basins of the headwaters of Ganga and its tributaries (Garhwal and Nepal) and Brahmaputra (Bhutan and Arunachal). Data compiled from refs 14, 23–32. The rocks from this region have highly radiogenic $^{87}Sr/^{86}Sr$.

cult to assess. Damming and impounding of rivers may reduce their water discharge, whereas the use of fertilizers may alter the natural Sr concentration of rivers and their $^{87}Sr/^{86}Sr$. We measured the Sr concentration and $^{87}Sr/^{86}Sr$ in a few fertilizer samples used in northern and western India. The results (Table 3) show that all the samples, except for the muriate of potash (MOP), have $^{87}Sr/^{86}Sr$ of ~0.708, consistent with that expected, as bulk of Sr in fertilizers is derived from limestones and phosphates of marine origin. The muriate of potash, which is abundant in potassium and rubidium, has a highly radiogenic $^{87}Sr/^{86}Sr$ of ~0.749 (Table 3). This high value raises concern about the impact such fertilizers can have in enhancing the natural $^{87}Sr/^{86}Sr$ of rivers. Our earlier observation[7] that $^{87}Sr/^{86}Sr$ for Ganga and many of its tributaries varies only marginally during different river stages is an indication that contamination, if any, of natural Sr by fertilizer may not be significant. Moreover, the most radiogenic Sr isotope composition is in the headwaters of Ganga and Ghaghara, where agricultural influences are minimal.

Summing up, our results on the Sr isotope systematics of rivers from India and Pakistan show that almost all the rivers sampled have $^{87}Sr/^{86}Sr$ in excess of the global average, with the headwaters of Ganga and its tributaries having the most radiogenic Sr isotope composition. Studies on the source of radiogenic Sr to the Ganga system (i.e. silicates or metamorphosed carbonates) would help in understanding better the coupling between Himalayan tectonics, weathering and climate change[22].

1. Burke, W. H., Denison, R. E., Hetherington, E. A., Koepnick, R. B., Nelson, H. F. and Otto, J. B., *Geology*, 1982, **10**, 515–519.
2. Palmer, M. R. and Edmond, J. M., *Earth Planet. Sci. Lett.*, 1989, **92**, 11–26.
3. Veizer, J., *Annu. Rev. Earth Planet Sci.*, 1989, **17**, 141–168.
4. Goldstein, S. J. and Jacobson, S. B., *Chem. Geol. (Isot. Geosci. Sect.)*, 1987, **66**, 245–272.
5. Sarin, M. M., Krishnaswami, S., Dilli, K., Somayajulu, B. L. K. and Moore, W. S., *Geochim. Cosmochim. Acta*, 1989, **53**, 997–1009.
6. Sarin, M. M., Krishnaswami, S., Somayajulu, B. L. K. and Moore, W. S., *Geochim. Cosmochim. Acta*, 1990, **54**, 1387–1396.
7. Krishnaswami, S., Trivedi, J. R., Sarin, M. M., Ramesh, R. and Sharma, K. K., *Earth Planet Sci. Lett.*, 1992, **109**, 243–253.
8. Sarin, M. M., Krishnaswami, S., Trivedi, J. R. and Sharma, K. K., *Proc. Indian Acad. Sci. (Earth Planet. Sci.)*, 1992, **101**, 89–98.
9. Pande, K., Sarin, M. M., Trivedi, J. R., Krishnaswami, S. and Sharma, K. K., *Chem. Geol.*, 1994, **116**, 245–259.
10. Krishnan, M. S., *Geology of India and Burma*, CBS Publishers and Distributors, India, 1982, 6th edn, pp. 536.
11. Pascoe, E. H., *A Manual of the Geology of India and Burma*, Government of India Press, Calcutta, 1963, vol. 3.

12. Valdiya, K. S., *Geology of the Kumaun Lesser Himalaya*, Wadia Institute Himalayan Geology, Dehra Dun, 1980, pp. 291.
13. Gansser, A., *Geology of the Himalayas*, Interscience, London, 1964, pp. 289.
14. Trivedi, J. R., Ph D Thesis, University of Gujarat, Ahmedabad, 1990, pp. 170.
15. Turekian, K. K. and Kulp, J. H., *Geochim. Cosmochim. Acta*, 1956, **10**, 245–296.
16. Edmond, J. M., *Science*, 1992, **258**, 1594–1597.
17. Palmer, M. R. and Edmond, J. M., *Geochim. Cosmochim. Acta*, 1992, **56**, 2099–2111.
18. France-Lanord, C., Derry, L. and Michard, A., in *Himalayan Tectonics* (eds Treloar, P. J. and Searle, M. P.), Geological Society Special Publication No. 74, pp. 603–621.
19. Dhoundial, D. P., Paul, D. K., Sarkar, A., Trivedi, J. R., Gopalan, K. and Potts, P. J., *Precambrian Res.*, 1987, **36**, 287–302.
20. Crawford, A. R., *J. Geol. Soc. India*, 1969, **10**, 117–167.
21. Naha, K., Srinivasan, R., Gopalan, K., Panthulu, G. V. C., Subba Rao, M. V., Vresbsky, A. B. and Bogomolov, V. E. S., *Proc. Indian Acad. Sci. (Earth Planet. Sci.)*, 1993, **102**, 547–565.
22. Raymo, M. E. and Ruddiman, W. F., *Nature*, 1992, **359**, 117–122.
23. Bhanot, V. B., Bhandari, A. K., Singh, V. P. and Kansal, A. K., *J. Geol. Soc. India*, 1979, **20**, 90–94.
24. Paul, D. K., Chandy, K. C., Bhalla, J. K., Prasad, R. and Sengupta, N. R., *Indian J. Earth Sci.*, 1982, **9**, 11–17.
25. Le Fort, P., Debon, F. and Sonet, J., *Geol. Bull. Univ. Peshawar*, 1979, 23–29.
26. France-Lanord, C. and Le Fort, P., *Trans R. Soc. Edinburgh*, 1988, **79**, 183–195.
27. Kunio, K., *Geochem. J.*, 1981, **15**, 63–68.
28. Deniel, C., Vidal, P., Fernandes, A. and Le Fort, P., *Contrib. Mineral Petrol.*, 1987, **96**, 78–92.
29. Xu, R. H., Scharer, U. and Allgre, C. J., *J. Geol.*, 1985, **93**, 41–57.
30. Sharma, K. K., Rao, D. R., Azmi, R. J., Gopalan, K. and Panthulu, G. V. C., *Curr. Sci.*, 1992, **62**, 528–530.
31. Honegger, K., Dietrich, V., Frank, W., Gansser, A., Thoni, M. and Tromunsdorff, V., *Earth Planet. Sci. Lett.*, 1982, **60**, 253–292.
32. Trivedi, J. R., Gopalan, K. and Valdiya, K. S., *J. Geol. Soc. India*, 1984, **25**, 641–654.

ACKNOWLEDGEMENTS. A number of our colleagues have helped in the sampling and field work. This include Prof B. L. K. Somayajulu, Dr Sheela Kusumgar, Dr N. Hussain, Dr K. K. Sharma, Mr Dhinakaran and Mr J. P. Bhavsar. Mr Ravi Bhushan and Mr Rengarajan helped with computation and graphics.

Received 29 March 1995; accepted 2 June 1995.

The non-induction of HSP70 in heat shocked Malpighian tubules of *Drosophila* larvae is not due to constitutive presence of HSP70 or HSC70

Bhupendra N. Singh and S. C. Lakhotia

Department of Zoology, Banaras Hindu University, Varanasi 221 005, India

To understand reasons for the earlier reported non-induction of any of the common heat shock polypeptides in Malpighian tubules of *Drosophila* larvae, we examined the levels of the 70 kDa heat shock polypeptide (HSP70) and of its cognates (HSC70) in control and heat shocked larval tissues, including the Malpighian tubules, by ^{35}S-methionine labelling in conjunction with Western blotting or by immunocytochemical localization using appropriate antibodies against the heat inducible and constitutively expressed members of the HSP70 family. The HSP70 was totally undetectable in western blots of control as well as heat shocked Malpighian tubules of *Drosophila* larvae; the levels of different HSC70 in untreated Malpighian tubules were comparable to those in other larval tissues and also did not increase following heat shock. Immunocytochemical localization of HSP70 showed that none of the highly polytenized cells of Malpighian tubules showed induction of HSP70 after heat shock. The non-induction of the common heat shock proteins in Malpighian tubules of *Drosophila* larvae, therefore, is not due to autoregulation by constitutively high levels of HSP70 or HSC70.

CELLS of virtually every organism respond to a sudden exposure to elevated temperature (heat shock) by transiently increased synthesis of a specific set of polypeptides, commonly referred to as heat shock proteins (HSPs) or stress proteins, which help cells survive the thermal and other damages[1–3]. This response, first discovered in *Drosophila*[4, 5], is a highly conserved biological phenomenon. However, contrary to the general dogma that all cells of an organism respond more or less uniformly to heat shock[1–3], Lakhotia and Singh[6], using ^{35}S-methionine to label the newly synthesized proteins, found that none of the well-characterized HSPs, including the most abundant 70 kDa heat shock protein (HSP70), was induced in heat shocked Malpighian tubules of *Drosophila* larvae. However, the basal levels of HSPs in the Malpighian tubules were not examined in that study[6]. Since synthesis of HSPs is autoregulated, particularly by HSP70 levels[2, 3, 7], the lack of induction of HSPs in heat-shocked Malpighian tubules[6] could be due to a constitutive presence of these proteins in threshold quantities so that further synthesis was inhibited. We have examined this possibility in the present study. Using ^{35}S-methionine labelling in conjuction with HSP70-specific antibodies, we reconfirm the earlier observation[6] on non-inducibility of the common HSPs in this tissue of *Drosophila* larvae and further show that the HSP70 is totally absent, before as well as after heat shock in this tissue; the levels of the various forms of 70 kDa

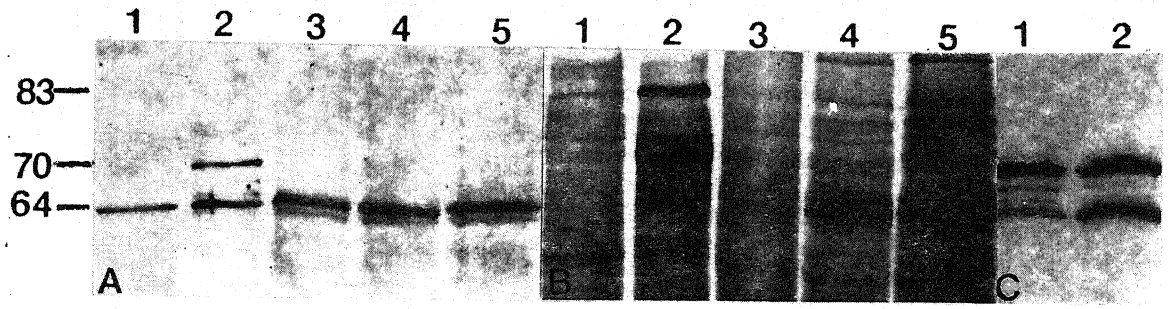


Figure 1. Showing absence of HSP70 in Malpighian tubules of *Drosophila* larvae. ***a***, Western blot of control (lanes 1 and 3) and heat-shocked (lanes 2, 4 and 5) samples of salivary glands (lanes 1 and 2) and Malpighian tubules (lanes 3–5) challenged with the HSP60 and HSP70 (*7Fb*) antibodies; lane 5 has 1.5 times more protein than lanes 3 and 4. ***b***, Fluorogram of the Western blot in (***a***). ***c***, Western blot of a sample of heat-shocked salivary glands alone (lane 1) or of a mixed sample of heat shocked salivary glands and Malpighian tubules challenged with HSP60 and HSP70 antibodies. For details, see text. The molecular sizes (in kDa) of polypeptides are indicated on sides.

heat shock cognate proteins (HSC70)[8, 9] in the larval Malpighian tubules are not higher than in other tissues. Therefore, the non-induction of HSP70 and other common HSPs in Malpighian tubules following heat shock is not due to inhibition by the constitutively present members of the HSP70 family.

A wild-type strain (Oregon R^+) of *Drosophila melanogaster* was used for these studies. The flies and larvae were reared on standard cornmeal–agar–yeast food at 22 ± 1°C.

Different organs (salivary glands, Malpighian tubules, brain ganglia) from late third-instar larvae were dissected in Poels' salt solution (PSS)[10], transferred to fresh PSS in microfuge tubes and were immediately either heat shocked at 37°C for 30 min or were kept at 22°C for 30 min as control. The heat shocked and control tissues were labelled with ^{35}S-methionine (BRIT, Bombay, 200 µCi/ml; sp.act. 1000 Ci/mM) for 30 min at 37°C and 22°C, respectively. Protein samples were prepared in the Laemmli sample buffer[11] and electrophoresed by SDS-PAGE as described[6]. Equal numbers of organs of a given type were taken for control and heat shock samples.

The electrophoresed polypeptides were transferred to nitrocellulose membrane (Schleicher & Schuell, Germany) by semi-dry blotting following the method of Lakhotia and Singh[12]. Two rat monoclonal antibodies, viz. *7Fb* (HSP70-specific) and *7.10* (recognizing the cognates (HSC70) as well as the heat inducible form of HSP70 of *Drosophila*)[13], obtained from S. Lindquist (USA), were used to detect the HSP70 and HSC70, respectively, in western blots at 1 : 1000 dilution. In some cases (Figure 1), the blots were first challenged with an anti-HSP60 antibody (*SPA-805*, Stressgen, Canada) and then with the HSP70-specific *7Fb* antibody. The primary antibodies were detected by using the corresponding HRP-conjugated secondary antibodies (Sigma, USA, or Bangalore Genei, Bangalore) at recommended dilutions. Following the immunodetection, the blots with ^{35}S-methionine labelled samples were fluorographed by dipping in toluene with 22% PPO for 3 min and drying for 1 h at room temperature before exposing to X-ray film to detect radiolabelled newly synthesized polypeptides. The Western blots and the fluorograms were then compared.

For immunocytochemical localization of HSP70, the salivary glands and Malpighian tubules with associated parts of the gut were dissected out from late third instar larvae and incubated in PSS at 22°C (control) or at 37°C (heat shock) for 30 min. Tissues were fixed in 4% paraformaldehyde in PBT (130 mM NaCl, 7 mM $Na_2PO_4 \cdot 2H_2O$, 3 mM $NaH_2PO_4 \cdot 2H_2O$, 0.2% Triton-X-100, pH 7.0) and processed for immunocytochemical localization of HSP70 using the *7Fb* antibody (1 : 200 dilution) and goat anti-rat secondary antibody (HRP conjugate, Sigma) at recommended dilution following the standard procedure[14].

Figure 1 *a* shows a Western blot of control and heat shocked salivary glands and Malpighian tubules from late third-instar larvae challenged sequentially with HSP60-specific (*SPA-805*) and HSP70-specific (*7Fb*) antibodies; fluorogram of the same blot is shown in Figure 1 *b*. Synthesis of all the HSPs was typically induced in salivary glands by heat shock (only HSP83, HSP70 and HSP68 are seen since only the upper half of the gels was always blotted). However, in agreement with the earlier finding of Lakhotia and Singh[6], heat shock to Malpighian tubules resulted in enhanced synthesis of some other polypeptides with a major band at 64 kDa (lanes 3–5 in Figure 1 *b*), but not of any of the usual HSPs (it may be mentioned here that this 64 kDa polypeptide was erroneously estimated as 58 kDa band by Lakhotia and Singh[6]). The *7Fb* antibody, which exclusively recognizes the heat shock inducible form of HSP70 of *Drosophila*[13], revealed the presence of HSP70 only in heat

shocked and not in control salivary glands. Neither the control nor the heat shocked Malpighian tubules showed any trace of HSP70 (lanes 3–5 in Figure 1). That the failure to get a signal with HSP70 antibody in Malpighian tubules was not due to the generally lower amount of proteins in Malpighian tubule samples was evident from the similar extent of labelling seen in the lanes for salivary glands and Malpighian tubules in the fluorogram in Figure 1 *b*. This was further confirmed by deliberately loading a greater amount of labelled sample from heat shocked Malpighian tubules in lane 5 (Figure 1 *a*, *b*); this also failed to reveal any trace of HSP70 in the Malpighian tubules. In addition, the same blot when challenged with an antibody for HSP60, which is constitutively present in all tissues of *Drosophila*[12, 15], showed more or less equal signal in all lanes, except in lane 5, which showed a more intense signal in proportion to the greater amount of protein loaded.

The complete absence of HSP70 in the Malpighian tubules can also not be due to a general failure of heat shock in these samples since the fluorogram clearly showed a significantly increased synthesis of the 64 kDa and other heat inducible polypeptides which are known to be heat shock induced in this tissue[6].

The HSP70 protein of *Drosophila* can decay *in vivo* as well as *in vitro* at a much faster rate: degradation, which is mediated by a proteolytic action of HSP70 protein upon itself, may sometimes occur rapidly, even during the course of electrophoresis[16]. To check that the absence of HSP70 in Malpighian tubules was not due to a tissue-specific degradation of this polypeptide, the Malpighian tubules and salivary glands excised from late third instar larvae were pooled, heat shocked together, dissolved in the sample buffer and run in one lane: Western blot revealed a distinct presence of the HSP70 in this mix (lane 2 in Figure 1 *c*), similar to that in the heat shocked sample of salivary glands alone (lane 1 in Figure 1 *c*). This demonstrated that the Malpighian tubule samples did not specifically lose their HSP70 due to degradation.

The HSP70 protein family in *Drosophila* includes both the heat-inducible (HSP70) and constitutively expressed homologues, the heat shock cognates (HSC70), which exist in several isoforms and show spatial and temporal tissue specificity[8, 17]. *Drosophila* embryos till the blastoderm stage do not synthesize HSP70 after heat shock, but are known to have an HSC70 in abundance[8]. Since it is possible that the total level of HSP70-related proteins in a cell is highly regulated[18], we examined whether the Malpighian tubules of *D. melanogaster* larvae contained higher levels of HSC70, which may prevent induction of HSP70. For this purpose, we used the *7.10* antibody, which recognizes[8, 13] both cognates and heat inducible forms of HSP70. Western blots of samples from different tissues (brain ganglia, salivary glands and Malpighian tubules of larvae), challenged with the *7.10* HSC70 monoclonal antibody, showed (Figure 2) a nearly similar

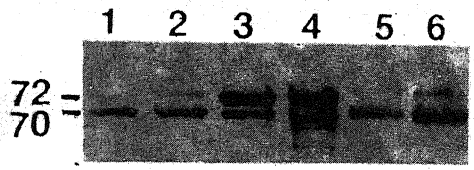


Figure 2. Levels of HSC70 in different larval tissues are similar. Western blot of control (lanes 1, 3 and 5) and heat shocked (lanes 2, 4 and 6) larval brain ganglia (lanes 1 and 2), salivary glands (lanes 3 and 4) and Malpighian tubules (lanes 5 and 6) challenged with the HSC70 antibody (*7.10*). The molecular sizes (in kDa) are indicated on the side.

pattern of reacting bands (68–72 kDa) in control and heat shocked samples of all the tissues. The more intense signal for salivary glands (lanes 3 and 4 in Figure 2) in this blot is due to more protein in these lanes. There was no noticeable increase in the levels of any of the HSC70 polypeptides in control or heat shocked Malpighian tubules (lanes 5 and 6 in Figure 2). This leads us to believe that the non-inducibility of HSP70 in Malpighian tubules is not due to a constitutive presence of high levels of HSC70.

Immunocytochemical staining with the HSP70-specific antibody (*7Fb*) confirmed the absence of HSP70 in Malpighian tubules since while the heat shocked salivary glands (Figure 3 *b*) and midgut (Figure 3 *d*) showed a distinct positive staining with this antibody, cells in control salivary glands (Figure 3 *a*), midgut (Figure 3 *c*) and in control as well as heat shocked Malpighian tubules (Figure 3 *c*, *d*) were negative for the HSP70 antibody. It is notable, however, that while the highly polytenized large cells of Malpighian tubules did not show any staining for HSP70, the fewer, smaller and flatter stellate cells[19] showed heat shock induced presence of HSP70 (see Figure 3 *e*). Apparently, the induced HSP70 in these cells was not enough to be detectable in the Western blots of whole Malpighian tubules. The strong nuclear signal in heat shocked salivary gland cells (Figure 3 *b*) is in agreement with the earlier report[20] that during heat shock the HSP70 concentrates strongly in nuclei.

It may be noted that while preparing the Malpighian tubule samples for electrophoresis, we took extreme care to exclude any part of the midgut to which these organs remain attached *in situ* (see Figure 3 *c*, *d*); we believe that the induction of all the common HSPs in Malpighian tubules of *Drosophila* larvae as reported in an earlier publication[5] was due to contaminating pieces of midgut, which like other tissues shows strong induction of HSP70 (Figure 3 *d*) and other HSPs.

The present results thus reconfirm the earlier report from our laboratory[6] that the Malpighian tubules of *Drosophila* larvae fail to synthesize any of the common

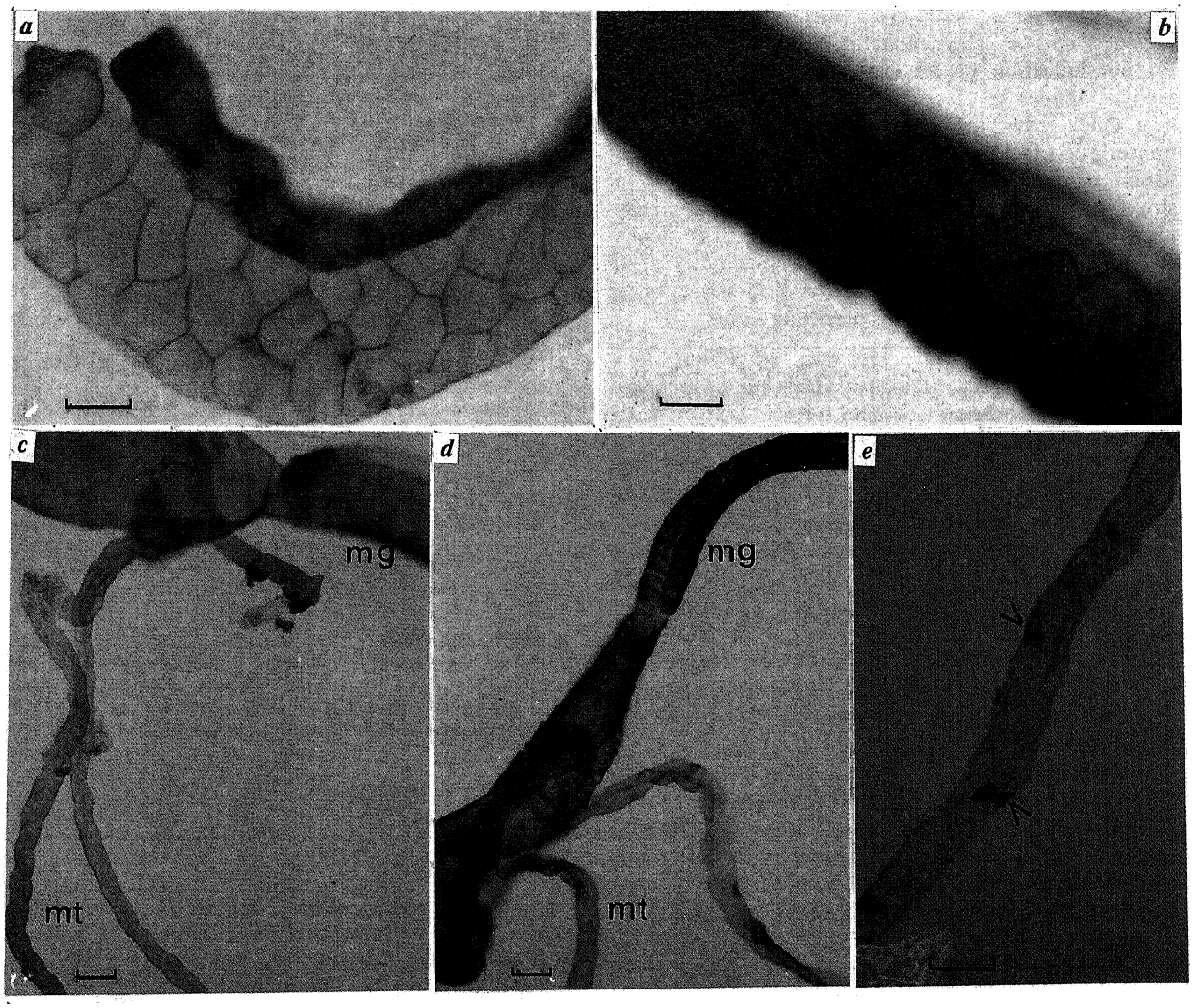


Figure 3. Immunocytochemical localization of HSP70 in control (*a* and *c*) and heat shocked (*b*, *d* and *e*) larval salivary glands (*a* and *b*), midgut (mg in *c* and *d*) and Malpighian tubules (mt in *c* and *d*). Part of a Malpighian tubule is shown at a higher magnification in (*e*) to reveal HSP70 induction in the peripherally located smaller stellate cells (arrowheads). Also note the stronger antibody binding in the nuclei of heat shocked salivary glands (*b*). The scale bars indicate 50 μm.

HSPs in response to heat shock. This study further shows that the non-induction of HSP70 in this tissue is not due to autoregulation[7, 18] by constitutively present HSP70 or HSC70. The same perhaps also applies to non-induction of the other HSPs in this tissue. The only other cell types of *Drosophila* which fail to mount the typical heat shock response are the nurse cells in late stages of oogenesis, oocyte and the pre-blastoderm embryos[21]. However, these cells, unlike the Malpighian tubules, do not show induction of a different set of polypeptides in response to heat shock, and also in these cases the non-induction of the HSPs may be correlated with a constitutively high level[8] of HSC70. Certain types of cells of *Sarcophaga*[22] also do not show synthesis of the typical HSP70 when heat shocked; however, in these cases a different form of HSP70, the HSP65, is induced instead. Unlike in *Sarcophaga*, the 64 kDa HSP induced in Malpighian tubules of *Drosophila* larvae is not a homologue of HSP70 but is a member of the HSP60 family[12, 15].

Several cell types in mammals also differ in inducibility of the HSP70 when heat shocked[23]; however, in most of these cases, unlike the situation in Malpighian tubules of *Drosophila* larvae, only the HSP70 is perhaps affected. Non-induction of the HSP70 in these mammalian cell types has been ascribed to transcriptional regulation by factors other than the heat shock transcription

factor[23]. Thus, compared to the above cases, the heat shock response of Malpighian tubules of *Drosophila* larvae is unique. The regulatory pathways responsible for non-induction of all the common HSPs in Malpighian tubules of *Drosophila* larvae may involve the heat shock transcription factor[24] and/or other auxiliary transcription factors[25] necessary for transcriptional activation of the heat shock genes under conditions of stress. Further studies will help understand the mechanism of this regulation as also the biological significance of this unique situation of non-inducibility of all the common HSPs and induction of a different set in the larger polytenized cells of this particular tissue.

1. Schlesinger, M. J., Ashburner, M. and Tissieres, A. (eds), *Heat Shock Proteins: From Bacteria to Man*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982.
2. Lindquist, S., *Annu. Rev. Biochem.*, 1986, **55**, 1151–1191.
3. Morimoto, R., Tissieres, A. and Georgopoulos, C. (eds), *Stress Proteins in Biology and Medicine*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1990.
4. Ritossa, F. M., *Experientia*, 1962, **18**, 571–573.
5. Tissieres, A., Mitchell, H. K. and Tracy, U., *J. Mol. Biol.*, 1974, **84**, 389–398.
6. Lakhotia, S. C. and Singh, A. K., *J. Genet.*, 1989, **68**, 129–137.
7. DiDomenico, B., Bugaisky, G. E. and Lindquist, S., *Cell*, 1982, **31**, 593–603.
8. Palter, B. K., Watanabe, M., Stinson, L., Mahowald, A. P. and Craig, E. A., *Mol. Cell. Biol.*, 1986, **6**, 1187–1203.
9. Lindquist, S. and Craig, E. A., *Annu. Rev. Genet.*, 1988, **163**, 631–677.
10. Lakhotia, S. C. and Mukherjee, T., *Chromosoma*, 1980, **81**, 125–136.
11. Laemmli, U. K., *Nature*, 1970, **227**, 680–685.
12. Lakhotia, S. C. and Singh, B. N., *Indian J. Exp. Biol.*, 1993, **31**, 301–302.
13. Kurtz, S., Rossi, J., Petko, L. and Lindquist, S., *Science*, 1986, **231**, 1154–1157.
14. Patel, N. H., *Methods Cell Biol.*, 1994, **44**, 445–488.
15. Singh, B. N. and Lakhotia, S. C. (in preparation)
16. Mitchell, H. K., Petersen, N. S. and Buzin, C. H., *Proc. Natl. Acad. Sci.*, 1985, **82**, 4969–4973.
17. Rubin, D. M., Mehta, A. D., Zhu, J., Shoham, S., Chen, X, Wells, Q. R. and Palter, K. B., *Gene*, 1993, **128**, 155–163.
18. Feder, J. H., Rossi, J. M., Solomon, J., Solomon, N. and Lindquist, S., *Genes Develop.*, 1992, **6**, 1402–1413.
19. Berridge, M. J. and Oschman, J. L., *Tissue Cell*, 1969, **1**, 247–272.
20. Velazquez, J. M. and Lindquist, S., *Cell*, 1984, **36**, 655–662.
21. Zimmermann, J. L., Petri, W. and Meselson, M., *Cell*, 1983, **32**, 1161–1170.
22. Joplin, K. H. and Denlinger, D. L., *J. Insect Physiol.*, 1990, **36**, 239–249.
23. Morimoto, R. I., Jurivich, D. A., Kroeger, P. E., Mathur, S. K., Murphy, S. P., Nakai, A., Sarge, K., Abravaya, K. and Sistonem, L. T., in *The Biology of Heat Shock Proteins and Molecular Chaperons* (eds Morimoto, R. I., Tissieres, A. and Georgopoulos, C.), Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, pp. 417–455.
24. Parker, C. S. and Topol, J., *Cell*, 1984, **37**, 273–283.
25. Lu, Q., Wallrath, L. L., Allan, B. D., Glaser, R. L., Lis, J. T. and Elgin, S. C. R., *J. Mol. Biol.*, 1992, **222**, 985–998.

ACKNOWLEDGEMENTS. This work was supported by a research grant from the Council of Scientific & Industrial Research, New Delhi, to SCL and by a research fellowship from Banaras Hindu University to BNS. We gratefully acknowledge the kind gift of *7.10* and *7Fb* monoclonal antibodies by Prof. S. Lindquist.

Received 21 April 1995; accepted 23 June 1995

Codon analysis of cyanobacterial genes

S. Krishnaswamy and S. Shanmugasundaram*

Bioinformatics Centre, Department of Biotechnology and *DBT Centre for BGA Biofertilizer, School of Biological Sciences, Madurai Kamaraj University, Madurai 625 021, India

The genes from different species of cyanobacteria have been analysed for relationships both between them and with other microbes and organelles. The codon bias indices indicate that the unicellular *Synechococcus* strain shows a preferred codon usage as those of low expressing genes of *E. coli*, *Agrobacterium tumefaciens* and *S. typhimurium*, whereas the filamentous *Anabaena* sp. and *Fremyella diplosiphon* strains show similarity to yeast (*Saccharomyces cerevisiae*), *Bacillus* sp., *B. subtilis* and *Streptococcus* sp. In so far as the codon bias index can be related to expression levels, the data suggest the possibility of suitable cloning hosts. The filamentous and unicellular cyanobacteria can be clearly separated based on the codon third-position bias and also the correlations involving codon usage. In general, correlations in codon usage profiles point to the filamentous strains being more related to gram positive bacteria and eukaryotes, whereas the unicellular strains are more related to gram negative bacteria.

PATTERNS of codon usage vary between organisms[1] and have helped to understand regulation of gene expression[2, 3]. Several organisms have been examined in detail in terms of their biased codon usage[4, 5]. However, only a preliminary analysis has been reported for codon usage in cyanobacterial genes[6] using a limited set of genes. The relationships within cyanobacterial species and with other organisms have been examined using the compilation by Wada *et al.*[7], who have tabulated the codon usage for different organisms. It must be mentioned that a study of codon usage of different genes from an organism, in spite of potential problems of the pattern of usage[8, 9], can lead to acceptable conclusions in certain cases and provide general trends[6, 7].

The codon usage data were taken from the compilation by Wada *et al.*[7]. These data refer to only complete genes and not to introns. The program CODON was used. This was developed by one of us earlier to examine the relationship of codon usage between plants and

Table 1. The percentage of synonymous codons for the unicellular (synechococcus, SYN) and filamentous (Anabaena, ANA); (Fremyella diplosiphon, FDI) strains are shown. The preferred codons selected with a cut off limit of sigma are marked with an asterisk

AA	cod	SYN	ANA	FDI
Gly	GGG	13	8	6
Gly	GGA	6	15	15
Gly	GGT	36*	53*	56*
Gly	GGC	45*	24	23
Glu	GAG	38*	21	17
Glu	GAA	62*	79*	83*
Asp	GAT	48*	58*	65*
Asp	GAC	52*	42*	35*
Val	GTG	34*	16	14
Val	GTA	7	32*	30*
Val	GTT	27*	36*	46*
Val	GTC	32*	16	10
Ala	GCG	22*	12	8
Ala	GCA	19*	27*	32
Ala	GCT	28*	42*	51*
Ala	GCC	31*	19	10
Arg	AGG	1	4	2
Arg	AGA	1	14	10
Arg	CGG	20	17	14
Arg	CGA	8	7	7
Arg	CGT	26	32*	34*
Arg	CGC	44*	26*	33*
Ser	AGT	13*	16*	14
Ser	AGC	24*	19*	28*
Ser	TCG	22*	5	3
Ser	TCA	6	15*	8
Ser	TCT	18*	28*	28*
Ser	TCC	18*	17*	19*
Lys	AAG	46*	31	40*
Lys	AAA	54*	69*	60*
Asn	AAT	28	44*	38*
Asn	AAC	72*	56*	62*
Ile	ATA	0	10	5
Ile	ATT	43	52*	51*
Ile	ATC	57*	38*	45*
Thr	ACG	21	5	2
Thr	ACA	6	32*	25
Thr	ACT	21	26*	31*
Thr	ACC	51*	37*	42*
Cys	TGT	33*	52*	38*
Cys	TGC	67*	48*	62*
Tyr	TAT	29	42*	35*
Tyr	TAC	71*	58*	65*
Leu	TTG	27*	26*	27*
Leu	TTA	7	25*	29*
Leu	CTG	36*	16*	13
Leu	CTA	6	15*	16
Leu	CTT	0	6	8
Leu	CTC	24	12	8
Phe	TTT	42*	57*	54*
Phe	TTC	58*	43*	46*
Gln	CAG	45*	27	26
Gln	CAA	55*	73*	74*
His	CAT	30	39*	58*
His	CAC	70*	61*	42*
Pro	CCG	34*	6	7
Pro	CCA	12	28*	21
Pro	CCT	18	37*	45*
Pro	CCC	35*	29*	28*

plant viruses[10]. The codon bias index, which determines the level of usage of preferred codons in a gene, has been defined earlier[4]. In this study, the same definition of the codon bias index is used. However, the selection of preferred codons is not based on the percentage occurrence of codons within a set of synonymous codons. The significance of a percentage occurrence depends on the number of synonymous codons. For example, 50% occurrence for a codon is not significant when there are only two synonymous codons but can be considered preferred if there are three or more synonymous codons in the set. Hence, we have calculated the standard deviation in percentage occurrence within the set of synonymous codons for each amino acid. Those codons in a synonymous set whose percentage occurrence is above a given cut-off limit times the standard deviation are flagged as 'preferred'. These preferred codons of the plant, for example, are used to determine if the codons of the plant virus are biased towards the plant codon usage. The best choice of the cut-off limit is one where the minimum number of codons is chosen as preferred while no amino acid is left without at least one preferred codon (Table 1).

In order to compare two codon usage profiles, for each amino acid a correlation coefficient is calculated using the percentage occurrence of codons within the synonymous set. The match coefficient for the entire set of amino acids is then defined as the sum of the weighted correlation coefficients. The weight for each amino acid is calculated as the ratio of the total number of codons for that amino acid to the total number of codons for all the amino acids of the two tables. This is done to take into account the differences in amino acid composition.

Thus,

$$\text{the match coefficient} = \sum_{i}^{N} w_i c_i,$$

where

N = all amino acids except Met, Trp,

w_i = weight factor,

defined as

$$w_i = \frac{\text{total number of codons for the } i\text{th amino acid}}{\text{total number of codons}},$$

and

c_i = correlation coefficient of synonymous codon usage,

$$c_i = \frac{\sum_{j}^{n_i} (p_{ji} - \bar{p}_i)(q_{ji} - \bar{q}_i)}{\left[\sum_{j}^{n_i} (p_{ji} - \bar{p}_i)^2\right]^{1/2} \left[\sum_{j}^{n_i} (q_{ji} - \bar{q}_i)^2\right]^{1/2}},$$

where P_{ji} is percentage occurrence of the jth codon of the ith amino acid of one table, q_{ji} is the percentage occurrence of the jth codon of the ith amino acid of the other table, $\bar{p}_i, \bar{q}_i$ are the means and n_i is the number of codons of the ith amino acid.

The match coefficient as defined here can vary from – 1 (for anticorrelated profiles) to zero (for profiles without correlation) to + 1 (for identical profiles). A high match coefficient indicates that amino acids with a higher frequency of occurrence have a similar distribution of codons within their set of synonymous codons.

The codon usage and the set of preferred codons[10] within the cyanobacterial species – the unicellular *Synechococcus* sp. (SYN), the filamentous *Anabaena* sp. (ANA) and *Femyella diplosiphon* (FDI) – are shown in Table 1. It is seen that the unicellular strain has a different usage from the filamentous strain, unlike in the earlier reported analysis[6]. This is also seen from the third-position codon bias (Table 2), where the unicellular strain has less A, T and more C, G as compared to filamentous strains. The match coefficients (Table 2) also show that there is a low correlation of codon occurrence patterns between unicellular and filamentous strains (0.31 in each case), while between the filamentous strains the coefficient is high (0.93). Thus, the match coefficient is able to provide discrimination between the two types of cyanobacterial strains. A comparison of the cyanobacterial codon usage profiles with those of other organisms using the match coefficients (Table 3) shows that the filamentous and unicellular strains of cyanobacteria also have different relationships with other organisms. This is seen from the unicellular strain having a high match coefficient with the gram negative bacteria *S. marcescens*, *Pseudomonas* sp., *A. vinelandii* and the gram positive *Streptomyces* sp. The filamentous strains have a high match coefficient with the gram positive bacteria *B. megaterium*, *Streptococcus* sp. and the eukaryotic yeast (*Schizosaccharomyces pombe* and *S. cerevisiae*), and with chloroplasts of chlamydomonas, pea, maize, spinach and tobacco. A high match coefficient shows a related pattern of codon occurrence. The above-mentioned difference in the codon usage profile implies that the unicellular and filamentous forms could have different evolutionary branches.

The codon bias index refers to how the preferred codons of other organisms are used in cyanobacterial genes, while the match coefficient is an indicator of the correlation in use of all the codons of cyanobacterial genes and the other organism. Thus, the bias and the match coefficient need not relate to each other as has also been shown elsewhere[10].

The usage of preferred codons in organisms has been related to levels of gene expression though it is known that there are other factors that also affect gene expression levels. A comparison of the codon bias indices of

Table 2. Differentiating unicellular from filamentous strain by codon analysis

	SYN 0	ANA –	FDI –
Frequency of base at position 3:			
A	0.15	0.29	0.26
T	0.23	0.31	0.35
C	0.36	0.24	0.25
G	0.27	0.17	0.14
Ratio of:			
XCG / XCC	0.73	0.33	0.26
XTA / XTT	0.29	0.78	0.72
Match coefficient relating the codon profile of the different forms of cyanobacteria:			
SYN 0	1.0	0.31	0.31
ANA –	0.31	1.0	0.93
FDI –	0.31	0.93	1.0

Table 3. The match coefficient relating the codon profile of the unicellular and filamentous forms with respect to those of the different species

Match coefficient	SYN 0	ANA –	FDT –
Gram negative bacteria			
E. coli	0.40	0.24	0.35
Pseudomonas sp.	0.60	–0.25	–0.24
A. tumefaciens	0.36	–0.29	–0.19
A. vinelandii	0.57	–0.25	–0.23
K. pneumoniae	0.55	0.11	–0.09
S. marcescens	0.75	0.04	0.05
S. typhimurium	0.43	0.17	0.26
R. capsulatus	0.50	–0.40	–0.31
B. japonicum	0.53	–0.36	–0.33
Rhizobium	0.54	–0.36	–0.35
Clostridium sp.	–0.38	0.47	0.37
Gram positive bacteria			
B. subtilis	0.04	0.42	0.47
Bacillus sp.	–0.08	0.46	0.47
B. megasterium	–0.08	0.66	0.67
B. sphaericus	–0.39	0.43	0.42
Streptococcus sp.	–0.08	0.62	0.62
S. aureus	–0.26	0.52	0.49
B. stearothermophilus	0.37	0.16	0.20
Streptomyces sp.	0.58	–0.34	–0.35
Eukaryotes			
S. pombe	–0.06	0.61	0.61
S. cerevisiae	–0.04	0.62	0.59
Kluyveromyces	–0.29	0.48	0.43
Wheat chloroplast	–0.22	0.60	0.57
Euglena gracilis	–0.28	0.54	0.51
Marchantia polymorpha	–0.25	0.55	0.52
Zea mays	–0.25	0.57	0.55
Pisum sativum	–0.23	0.61	0.60
Spinach	–0.25	0.58	0.57
Nicotiana tabaccum	–0.18	0.62	0.61
Chlamydomonas	0.10	0.73	0.63

Table 4. The codon bias indices (CBI) for the unicellular and filamentous forms with respect to the preferred codons selected for the different species. The cut-off limit used for selection in standard deviation units is given in parantheses

CBI	SYN 0	ANA –	FDI –
Gram negative bacteria			
E. coli (low) (2)*	0.31	0.10	0.14
E. coli (high) (2)**	0.17	0.17	0.21
E. coli (2)	0.34	0.20	0.22
Pseudomonas sp. (2)	0.22	–0.09	–0.09
A. tumefaciens (2)	0.28	–0.06	–0.04
A. vinelandii (2)	0.21	–0.06	–0.04
K. pneumoniae (2)	0.21	–0.07	–0.09
S. marcescens (2)	0.21	–0.05	–0.05
S. typhimurium (2)	0.26	–0.01	–0.07
R. capsulatus (2)	0.20	–0.10	–0.12
B. japonicum (2)	0.22	–0.13	–0.13
Rhizobium (2)	0.22	–0.14	–0.15
Clostridium sp. (2)	–0.18	0.20	0.22
Gram positive bacteria			
B. subtilis (2)	0.09	0.28	0.29
Bacillus sp. (2.5)	0.13	0.30	0.29
B. megaterium (2.5)	–0.03	0.20	0.17
B. sphaerieus (2.0)	–0.19	0.22	0.23
Streptococcus sp. (2.0)	–0.05	0.27	0.33
S. aureus (2.0)	–0.12	0.24	0.26
B. stearothermophilus (2.5)	0.16	0.00	–0.04
Streptomyces sp. (2.0)	0.17	–0.08	–0.07
Eukaryotes			
S. pombe (2)	–0.03	0.20	0.26
S. cerevisiae (2)	0.01	0.28	0.32
Kluyveromyces (2)	–0.12	0.21	0.23
Wheat chloroplast (2)	–0.17	0.25	0.30
M. phymorpha (2)	–0.04	0.19	0.22
Z. mays (2)	–0.17	0.26	0.29
P. sativum (2)	–0.14	0.24	0.26
Spinach (2)	–0.22	0.24	0.27
N. tabaccum (2)	–0.09	0.26	0.29
Chlamydomonas (2)	–0.02	0.27	0.28

*Genes with low expression and **genes with high expression.

the cyanobacterial genes with respect to other organisms (Table 4) shows that the codon usage of unicellular genes is biased towards *E. coli* (low expression) and *A. tumefaciens*, whereas the filamentous genes are biased towards *S. cerevisiae*, *Bacillus* sp., *B. subtilis* and chloroplasts. In as far as the biased codon usage is a determinant of the levels of gene expression, it is seen through this analysis that *A. tumefaciens* could form an alternate cloning host apart from *E. coli* for unicellular strains and filamentous strains may be better expressed in *S. cerevisiae*, *Bacillus* sp. or *B. subtilis*. These conclusions are being made to form a basis for molecular biology experiments involving cyanobacteria.

1. Andersson, G. E. and Kurland, C. G., *Microbiol. Rev.*, 1990, **54**, 198–210.
2. Ikemura, T., *J. Mol. Biol.*, 1981, **146**, 1–21.
3. de Boer, H. A. and Kastelein, R. A., in *Biased Codon Usage: An Exploration of its Role in Optimization of Translation* (eds Reznikoff, W. and Gold, L.), Buttersworth, Boston, 1986, pp. 225–286.
4. Bennetzen, I. L. and Hall, B. D., *J. Biol. Chem.*, 1982, **257**, 3026–3031.
5. Sharp, P. M. and Devine, K. M., *Nucleic Acids Res.*, 1989, **17**, 5029–5039.
6. Tandeau de Marsac, N. and Houmard. J., in *The Cyanobacteria* (eds Fay, P. and Van Baalen, C.), Elsevier, 1987, pp. 251–302.
7. Wada, K., Wada, Y., Doi, H., Ishibashi, F., Gojobori, T. and Ikemura, T., *Nucleic Acids Res.*, 1992, **20** (Suppl.), 2111–2118.
8. Lloyd, A. J. and Sharp, P. M., *Mol. Gen. Genet.*, 1991, **230**, 288–294.
9. Sharp, P. M., Cowe, E., Higgins, D. G., Shields, D. C., Wolfe, K. H. and Wright, F., *Nucleic Acid Res.*, 1988, **16**, 8207–8210.
10. Sudha, S. N., Krishnaswamy, S. and Vaithilingam Sekar, *Curr. Sci.*, 1992, **63**, 573–575.

ACKNOWLEDGEMENTS. We are thankful to DBT, New Delhi, for financial support to Bioinformatics Centre, MKU. The work was done under the Technology Development and Demonstration Project on blue–green algae funded by DBT, New Delhi.

Received 24 July 1993; revised accepted 23 April 1995

Effects of feeding mulberry (*Morus* sp.) leaves supplemented with different nutrients to silkworm (*Bombyx mori* L.)

A. A. Sarker, M. R. Haque, M. A. Rab and N. Absar*

[1]Bangladesh Sericulture Research and Training Institute, Rajshahi 6100, Bangladesh
*Department of Biochemistry, Rajshahi University, Rajshahi 6205, Bangladesh

The growth of silkworm larvae improved significantly on feeding them mulberry leaves supplemented with different nutrients. The total protein content of silk gland increased, while the water, total lipid and carbohydrate contents decreased. The larval weight, silk gland weight and total protein in silk gland were found to be highest for treatment TS_8 (soymilk + sugar + vitamins + potassium iodide salt), followed by TS_7 (glycine + alanine + sugar + vitamins + potassium iodide salt) and TS_9 (milk powder + sugar + vitamins + potassium iodide salt) treatments. The highest values of single-cocoon weight, single-shell weight and filament length were observed for treatment TS_8 (1.34 g, 0.26 g and 799.42 m/cocoon, respectively), whereas cocoon yield was found to be highest for treatment TS_7 (110.17 g/100 larvae). No significant difference was observed between the treatments TS_7, TS_8 and TS_9, their influence on silk gland and cocoon characteristics being more or less similar. The cocoon yield/100 larvae was increased by 20.61%, 20.38% and 19.06% in treatments TS_7, TS_8 and TS_9, respectively, over the control treatment.

IT has been well established that the silkworm (*Bombyx mori* L.) requires certain essential sugars, proteins, amino acids, fatty acids and vitamins for its normal growth, survival and also for the growth of silk gland and higher production of good-quality silk. The nutritional status of the mulberry leaves can be improved by enriching them with extra nutrients. Fortification of mulberry leaves with extra nutrients like glucose, glycine, egg albumin, molasses, etc., was found to increase the larval growth and improve cocoon characteristics[1, 2].

Feeding the silkworms mulberry leaves fortified with 2% solution of casein or jaggery or urea in tap water did not improve any of the economic cocoon characteristics significantly[3]. Gridhar and Radha[4] reported that supplementation with 10 ppm glycine resulted in higher larval weight, shell weight, filament length compared to control. Alagumalai *et al.*[5] observed that fortification of mulberry leaves with the flours of black gram and red gram bran improved the larval growth and cocoon characteristics. Young larvae reared on 1.5% ascorbic acid enriched mulberry leaves resulted in higher silk length, weight and values[6].

The work was undertaken to study the biochemical composition of silk gland as well as the growth and cocoon production of silkworm *Bombyx mori* L. on feeding mulberry leaves supplemented with different nutrients.

The silkworm race BSRI-85/3 was selected for the study. Normal mulberry leaves were supplemented with different nutrient chemicals (variety BM-3) in the form of aqueous solution as follows:

TS_0 (control): Without supplement
TS_1 : Glycine (0.5%) + alanine (0.5%)
TS_2 : Soymilk
TS_3 : Milk powder (10%)
TS_4 : Sugar (10%)
TS_5 : Vitamin B (0.5%) + vitamin C (1%)
TS_6 : Potassium iodide salt (1%)
TS_7 : $TS_1 + TS_4 + TS_5 + TS_6$
TS_8 : $TS_2 + TS_4 + TS_5 + TS_6$
TS_9 : $TS_3 + TS_4 + TS_5 + TS_6$

Mulberry leaves supplemented with these nutrients in stated concentrations were fed to silkworms from fourth instar of larval stage till spinning since during this period the worms have a more vigorous growth with larger assimilation of nutrients. Worms were reared together in mass beds up to the second larval stage and after second-moult worms were separated in four batches (replication) of 100 worms for each treatment. Worms were fed with mulberry leaves smeared with aqueous solutions of above nutrients just before feeding, at the rate of 5–10 ml solution per 30–50 g mulberry leaves.

The weights of mature larvae were determined by taking the weight of 10 larvae just before spinning, and the other individual quantitative and qualitative cocoon characteristics were determined after the spinning following standard techniques as applied for the assessment of silk cocoons.

Silk gland was isolated in the fifth instar of larval stage. The fourth batch (replication) was used for this experiment. The fresh weight and dry weight of silk gland were estimated gravimetrically. The carbohydrate content[7], total lipids[8] and total proteins[9] were estimated.

Soymilk was prepared according to the methods of Khaleque *et al.*[10] and Sarker and Khaleque[11]. The rearing was conducted (in 1992) following the method of Krishnaswami[12] during the seasons S_1 (February–March), S_2 (May–June), S_3 (August–September) and S_4 (November–December). The temperature and relative humidity of the rearing room were, respectively, 27.30 ± 3.50°C and 67.60 ± 4.90%, 29.60 ± 3.70°C and 73.50 ± 7.30%, 28.40 ± 0.60°C and 84.30 ± 5.20%, 23.65 ± 3.20°C and 69.30 ± 4.10% in different seasons.

The chemical composition of silk gland and the growth of *Bombyx mori* larvae fed mulberry leaves supplemented with different nutrients were studied in four seasons. Table 1 presents the mean larval weight, silk gland weight and the chemical composition of silk gland under different treatments of supplementation over the four seasons. The highest larval weight (3.62 g/individual) and silk gland weight (778.75 mg/larva)

Table 1. Mean weight and chemical composition of silk gland fed on mulberry leaves supplemented with different nutrients

Treatment	Average larval weight (g)	Silk gland weight (mg/larva)	Chemical component of silk gland			
			Water (mg/g wt)	Total protein (mg/g dry wt)	Total lipid (mg/g dry wt)	Carbohydrate (mg/g dry wt)
TS_0	2.89e	572.50d	736.85a	269.73d	634.35a	6.02a
TS_1	3.30bc	695.00bc	719.26bc	313.13b	620.22abc	5.83ab
TS_2	3.37b	708.25b	718.17bc	314.71b	615.00bcde	5.81ab
TS_3	3.27bcd	681.25bcd	720.13bc	313.24b	616.01bcd	5.77ab
TS_4	3.09de	656.25bcd	725.45bc	282.34c	617.16bcd	5.84ab
TS_5	3.02e	649.50bcd	724.78bc	287.19c	621.22ab	5.72ab
TS_6	2.99e	627.50cd	726.42b	286.03c	619.41bc	5.71ab
TS_7	3.60a	760.75a	704.38d	342.20a	604.13de	5.42b
TS_8	3.62a	778.75a	700.48d	343.21a	606.27cde	5.39b
TS_9	3.60a	741.50a	701.08d	340.87a	601.82e	5.35b

Figure(s) followed by the same letter(s) are not significantly different at 1% level as per Duncan's new multiple range test.

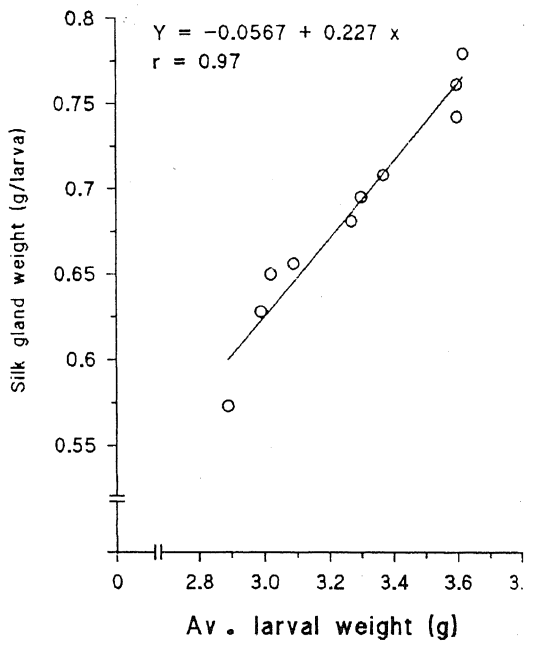


Figure 1. Mean larval weight as a function of silk gland weights.

were observed on feeding mulberry leaves subjected to treatment TS_8, followed by TS_7 (3.60 g and 760.75 mg/larva, respectively) and TS_9 (3.60 g and 741.50 mg/larva, respectively). The protein content of silk gland increased while the water, total lipid and carbohydrate contents of silk gland decreased significantly due to supplementation of mulberry leaves with different nutrients. The highest total protein (343.21 mg/g) content of silk gland was noticed on feeding mulberry leaves subjected to treatment TS_8, while the water (736.85 mg/g), total lipid (634.35 mg/g) and carbohydrate (6.02 mg/g) contents silk gland were found to be highest in the control treatment (TS_0). It may be seen from the Table 1 that the values for larval weight, silk gland weight and chemical components of silk gland on feeding mulberry leaves subjected to treatments TS_7, TS_8 and TS_9 are more or less similar and did not vary significantly.

The mean larval weight was plotted against silk gland weight and the regression line was fitted by the method of least squares (Figure 1). From statistical analysis the correlation coefficient was found to be positive ($r = 0.97$), which indicates that as the larval weight increases the silk gland weight increases.

Mean cocoon characteristics under different treatments of supplementation over four seasons are presented in Table 2. It is seen from the table that single-cocoon weight, single-shell weight, cocoon yield and filament length increased significantly by feeding mulberry leaves supplemented with different nutrients. All the cocoon characteristics were found to show the

Table 2. Mean cocoon characteristics of BSRI-85/3 race fed on mulberry leaves supplemented with different nutrients

Treatment	Single-cocoon weight (g)	Single-shell weight (g)	Cocoon yield/100 larvae (g)	Filament length (m/cocoon)
TS_0	1.27f	0.19e	91.34f	684.00h
TS_1	1.30cde	0.22cd	103.16bcd	754.16cde
TS_2	1.31bcd	0.23bc	104.79b	758.50bcd
TS_3	1.30cde	0.22cd	103.45bc	745.75de
TS_4	1.28ef	0.21cde	99.99bcd	724.08f
TS_5	1.27f	0.20de	97.31de	713.08fg
TS_6	1.27f	0.20de	94.92def	700.58gh
TS_7	1.33ab	0.25ab	110.17a	776.00ab
TS_8	1.34a	0.26a	109.96a	779.42a
TS_9	1.32abc	0.25ab	108.75a	771.58abc

Figure(s) followed by the same letter(s) are not significantly different at 1% level as per Duncan's new multiple range test.

maximum improvement on feeding mulberry leaves subjected to treatment TS_8 except for cocoon yield, which was found to be highest for treatment TS_7. The lower cocoon yield in treatment TS_8 might be due to the higher mortality. However, no significant differences were observed between the treatments TS_7, TS_8 and TS_9 in respect of the cocoon characteristics of silkworm *Bombyx mori* L.

Supplementation with amino acids like tyrosine, phenylalanine and alanine improved the leaf quality of mulberry[13]. Sengupta *et al.*[1] reported that the silk content was more with 1% ascorbic acid in the diet of the silkworms. Babu *et al.*[6] observed that the first and second instar larvae reared on 1.5% ascorbic acid enriched mulberry leaves resulted in higher silk length, weight and denier values. Salts of potassium and calcium have been found to have significant effect on the growth of the silkworm. The addition of amino acids in low concentrations gave the best results regarding weights of larvae, percentage of silk gland, crude protein, cocoons and cocoon shells[14]. Beneficial effect of feeding sugars to larvae of the *Bombyx mori* silkworm was noted by Ito[15]. Fortification of mulberry leaves with the flours of black gram bran and red gram bran increased the larval weight, silk gland weight and improved cocoon characteristics like cocoon weight, shell length, cocoon length,cocoon width, shell ratio and also the reproductive parameters[5]. Observations made during this experiment corroborates these findings to a great extent. The supplementation with glycine + alanine + sugar + potassium iodide salt + vitamins (treatment TS_7), soymilk + sugar + potassium iodide salt + vitamins (treatment TS_8) and milk powder + sugar + potassium iodide salt + vitamins (treatment TS_9) significantly improved the larval weight, silk gland weight, chemical composition of silk gland and cocoon characteristics of *Bombyx mori* (Tables 1 and 2). Cocoon yield/100 larvae was increased by 20.61%, 20.38% and 19.06% with treatments TS_7, TS_8 and TS_9, respectively, over the

control (no supplementation). These might be due to the presence of increased amounts of proteins/amino acids and other nutrients in mulberry leaves which contributed to the extra growth of the silkworm, thereby increasing the weights of larvae and silk gland and improving the cocoon characteristics of *Bombyx mori*.

There existed a positive correlation ($r = 0.77$) between silk gland weight and larval weight in the fifth instar larvae fed on mulberry leaves supplemented with different nutrients. Opender *et al.*[16] and Devaiah *et al.*[17] observed that the rate of increase in silk gland weight is proportional to the rate of increase in the body weight of silkworm. It has previously been reported that the ratio of the two rates was constant throughout the fifth instar[18]. The present study also confirms these results.

The significant decrease in the water, total lipid and carbohydrate contents and the increase in the total protein content in silk gland indicate the qualitative and quantitative improvement of silk gland due to feeding mulberry leaves supplemented with different nutrients, subsequently reflected in the improved cocoon characteristics of *Bombyx mori* L. However, further studies are needed on large-scale level before transfer of technology to the field.

1. Sengupta, K., Singh, B. D. and Mustafi, J. C., *Indian J. Seric.*, 1972, **11**(1), 11–19.
2. Verma, A. N. and Atwal, A. S., *Indian J. Seric.*, 1963, **2**(4), 1–4.
3. Kumararaj, S., Vijayaraghavan, K. and Krishnaswami, S., *Indian J. Seric.*, 1972, **11**(1), 68–72.
4. Gridhar, P. and Radha, N. V., National Seminar on prospects and problem of sericulture in India, 27–30 March, 1986, p. 18.
5. Alagumalai, K., Shermila Hepshyba, C. S. and Ramaraj, P., *Indian Silk*, 1991, **30**(6), 10–11.
6. Babu, M., Swamy, M. T., Rao, P. K. and Rao, M. S., *Indian J. Seric.*, 1992, **31**(2), 111–114.
7. Carrol, N. V., Longle, R. W. and Roe, J. H., *J. Biol. Chem.*, 1956, **220**, 583–593.
8. Folch, J. M., and Lees, M. P., Stane-stanely, G. H., *J. Biol. Chem.*, 1957, **226**, 497–509.
9. Lowry, O. H., Roserbrough, N. J., Farr, A. L. and Randall, R. J., *J. Biol. Chem.*, 1951, **193**, 265–275.
10. Khaleque, A., Bannatyne, W. R. and Walace, G. M., *J. Sci. Fd. Agric.*, 1970, **21**, 579.
11. Sarker, A. A. and Khaleque, A., *Bangladesh J. Agric.*, 1978, **3**, 227–282.
12. Krishnaswami, S., Bulletin No. 2, CSRTI, Mysore, 1987.
13. Nagarajan, P. and Radha, N. V., *Indian Silk*, 1990, **29**(4), 21–22.
14. Moustafa, S. M. and Elkaraksy, L. A., *Agric. Res. Rev.*, 1990, **66**(1), 105–109.
15. Ito, T., *J. Insect Physiol.*, 1960, **5**, 95–107.
16. Opender, K., Tikku, K., Saxena, B. P. and Atal, C. K., *Indian J. Seric.*, 1979, **18**, 1–5.
17. Devaiah, M. C., Rajashekhargouda, R., Suhas, Y. and Govinadan, R., *Indian J. Seric.*, 1985, **24**(1), 33–35.
18. Ueda, S., Kimura, R. and Suzuki, K., *Bull. Seric. Exp. Sta. (Tokyo)*, 1971, **25**, 1.

Received 15 November 1994; revised accepted 23 May 1995

The Guide to the Galaxy. N. Henbest and H. Couper. Cambridge University Press, Cambridge. 1994. pp. 265. Price: £35/$49.95 (hardbound); £17.95/$24.95 (paperback).

The last two decades have seen great advances in astronomical colour photography and in computer-generated images of data obtained at all wavelengths from both ground-based and space observatories. These made possible the appearance of 'picture books' in astronomy, of which there have been several in recent years. Apart from adorning coffee tables, which they do very well, such publications serve to convey to nonastronomers some idea of the great beauty there is in the heavens albeit inaccessible to the unaided eye. They also transmit some of the excitement surely felt by the astronomers in the difficult and time-consuming process of creating such images. Even more useful is the role that such books can play in arousing the interest of the young, if not in pursuing a career in astronomy, at least to learn of some of its wonders. The book under review has gone much further and could well become a Bible for the amateur astronomer who really wants to know what he or she is looking at whether through a small telescope, or binoculars, or just gazing at the night sky.

A short description of the book on the flyleaf reads:

'Nigel Henbest and Heather Couper take us on a personal tour of the Milky Way in this delightful guide to our home Galaxy. This book is written in non-technical language, for the general reader and the amateur astronomer. The lucid text is supported by a dazzling portfolio of colour photographs of stars, star clusters and gas clouds in the Milky Way. Specially created maps locate the hundreds of tourist spots visited in this journey through space.

The guide introduces the nature and structure of our Galaxy as a whole. All types of objects that it contains are then described. These include mysterious black holes, the giant globular clusters of stars, every kind of star, regions of star formation, exploding stars and their remnants and the enigmatic objects at the galactic centre, as well as the giant halo surrounding the Galaxy.'

That is accurate as far as it goes, but there is more to the book than that. On the very first page the authors call their effort 'a celebration of the first clear, panspectral view of our own star-city . . ., a book that could not have been written until now'. Putting together what we have seen through all the astronomical windows from radio to gamma rays, the authors succeed in presenting a 'holistic' picture of our Galaxy illustrated with maps that really 'could not have been drawn up even a few years ago'.

Scattered all through the book, and particularly in the early chapters, are engaging accounts of how certain profound discoveries and advances came to be made by the great names in the history of astronomy. A score of portraits/photographs of particular individuals starting with Galileo adds to the character of the book, much credit for which must surely go to Michael Marten, the picture editor.

The word 'Guide' in the title of the book is serious. The book is aimed at the genuinely interested visitor to the Galaxy and not the kind of tourist that would like to do the Taj Mahal in ten minutes and get back to the hotel. Adopting a geographical approach the authors lead the reader firmly by the hand through the highways and byways of our Milky Way system and keep feeding more information than any intelligent lay reader could ever absorb in one go. The maps, which are brightly coloured, are innovations and contain a lot of information obtained from a variety of sources, much of it in very recent years. There are a great number of large boxes which I found somewhat distracting at first; in retrospect, however, I feel that the number of topics covered could not have been handled otherwise. Indeed, I find it amazing how much astrophysical knowledge has been touched upon as and when relevant to whatever object in the Galaxy or phenomenon in the sky was being described.

Everything said above could well be called praise. Is there any criticism I have of this fine book? Almost nothing as far as the content is concerned, but there were a few places where the choice of words grated a bit. An example of such infelicity is (p. 125): 'Determined to punish Orion's chutzpah, Hera . . . . . '. Something else which appeared more than once is, . . . star X or black hole Y or whatever weighing '*in at*' so many Suns, pp. 205, 245, etc. Astronomical objects are not boxers or jockeys and the authors could please remove the prepositions in future editions.

To sum up, *The Guide to the Galaxy* is a must for the serious amateur, a worthwhile acquisition for any scientific library and even for the professional astronomer, a source of information outside one's own area of specialization.

V. RADHAKRISHNAN

Raman Research Institute, Bangalore 560 080, India

Crustal Evolution of Singhbhum – North Orissa, Eastern India. A. K. Saha. Memoir 27, Geological Society of India, Bangalore. 1994. pp. xii + 341. Price: Rs. 350.00; $50.

The Precambrian terrain of Bihar–Orissa has attracted the attention of Indian geologists for more than a century, probably because of the presence of rich mineral deposits. The work of Dunn[1,2], Jones[3], Krishnan[4] and Dunn and Dey[5] laid down the basic framework of the geology of this region. Since the middle part of this century, there has been a spurt of research activity on structural, petrological, geochemical and geochronological aspects and the new work has greatly enriched our knowledge on this important segment of the Indian shield. Saha and his group have been in the forefront of this modern research and several of our current notions stem from their work. Saha has drawn on his vast experience in writing this monograph on the geology of Singhbhum–North Orissa, which would remain a standard work of reference for many years to come. In a book of this type the author's own perceptions are bound to come out strongly in the writing, though Saha has made conscious effort to refer to opposing viewpoints. As the title of the book indicates, he has attempted to integrate the information coming from different lines of study to build a model of Precambrian crustal evolution of this terrain.

Chapter 1 places the Singhbhum region in the perspective of the Indian shield by highlighting the salient points about the evolutionary history of the different Precambrian provinces. No mention is, however, made about the Precambrian sedimentary basins, e.g.

the Vindhyans and the Cuddapahs. A brief summary of the current status of our knowledge on the Singhbhum–North Orissa region in Chapter 2 is followed by the chapter describing the regional structural framework of the terrain. The work of R. K. Verma and his coworkers on the gravity field is summarized. Upper crustal structures interpreted from the gravity data are described. The thermal model for the formation of the West Singhbhum Iron Ore Basin is discussed in some detail. However, one of the major tectonic problems of this region, namely the structural relations between the North Singhbhum fold belt and the Singhbhum Granite Complex and the Iron Ore Basins, is not thoroughly discussed. It is stated that the Archaean nucleus subducted southward below the Granulite Belt, but no firm supporting evidence is offered, nor is there any attempt to explore how this southward subduction fits with the supposed northward subduction along the northern boundary. A convincing demonstration of the subduction of an Archaean (>3.2 Ga) cratonic nucleus below another plate during the Proterozoic would have been of great geodynamic interest.

Chapters 4–8 describe the different units in the Archaean craton: Older Meamorphic Group and Older Metamorphic Tonalitic Gneiss, Iron Ore Group, Singhbhum Granite, Bonai Granite, Nilgiri Granite, and Pallahara Gneiss. Saha and his group have carried out sustained work on these rocks and the chapters contain a wealth of information, particularly on field relations, petrography and geochemistry. The structural geometries in the different units are described and the evidences for inferring the relative ages of the supracrustal and granitoid rocks are cogently discussed. Modal and chemical (major, trace and REE) analyses of the different rock types provide valuable documentation. Interesting new results on the petrogenesis of the granitic rocks are presented. These are based on petrogenetic modelling from geochemical data, a novel exercise in Indian Precambrian geology. Such studies provide important constraints in building up tectonic models.

Chapter 9 summarizes the available data on the geochronology of these rocks. The author has emphasized the importance of multiple methods of dating. However, for a region of such complexity the dates are too few and the picture that emerges is rather confusing. Some of the dates are interpreted as metamorphic dates, others as ages of crystallization while still others are thought to represent partially reset ages. However, no clear rationale for such deductions are presented. It is disquieting that Saha is adhering to the idea of 3.8 Ga old OMTG despite the fact that Basu and his coworkers could not reproduce their earlier isochron. In a recent publication Sharma *et al.*[6], have repudiated their earlier conclusion and have upheld the *ca.* 3.3 Ga age of the OMTG. It is fairly certain that the OMTG is at least 3.3 Ga old and during the next one billion years there were several episodes of granitic activity. But much more work needs to be done to discriminate geochronologically all the different events: the 3.8 Ga emplacement age of OMTG is still to be established convincingly.

Chapters 11–14 deal with the Proterozoic fold belt of North Singhbhum. The broad structural features and metamorphic pattern are described, though no reference is made to some new contributions on structure and metamorphism in the belt, for example, Bhattacharya[7], Bhattacharya *et al.*[8], Chakraborty and Sen[9], Mukhopadhyay *et al.*[10], and Roy[11]. The Dalma volcanic belt is described as a synclinorium though there is no structural evidence in favour of this idea nor are there any data to suggest that the supracrustal sequences north and south of the volcanic belt are equivalent. A welcome inclusion is the sedimentary facies analysis of a part of the Chaibasa Formation. There is an excellent treatment on the geochemistry of the volcanic rocks, particularly the Dalmas and the Dhanjoris. The compilation of the chemical parameters of the different metabasic suites, the granitoid rocks and the other intrusive rocks would be useful for all earth scientists interested in this belt. Petrogenesis of the Mayurbhanj Granite has been discussed in some detail using mass balance calculations.

Chapter 17 describes the major thrust zones in the region, the most important of which is the Singhbhum shear zone. The author has preferred to use the term Copper Belt Thrust, though it is a ductile shear zone several kilometers wide and no discrete thrust surface can be mapped. The thrust is stated to grade into a high-angle gravity fault in the west (p. 177), but there is no structural indication of gravity faulting. There are scanty structural data on the Northern Shear Zone and even less on the Sukinda Thrust Zone.

The author's synthesis on the Precambrian crustal evolution of the Singhbhum–North Orissa region is summarized in Chapter 19. It is implicit from his discussion that the model is not tightly constrained. There are many gaps in our knowledge, some of which the author has highlighted; these need to be bridged before we arrive at a fuller understanding of this part of the Indian Shield.

A large number of well-drafted maps and diagrams supplementing the text have added to the value of this monograph. Unfortunately, many of the localities referred to are not marked on the geological map (Figure 2.3) nor are their latitudes–longitudes given. Some mistakes have been noticed in the list of references.

The monograph presents a fairly balanced view of the existing state of our knowledge on the different aspects of the geology of this terrain. The wealth of petrological and geochemical data culled from different sources and the exhaustive bibliography are the main attractions of this volume. It should find a place in the bookshelf of all serious students of Precambrian geology.

1. Dunn, J. A., *Mem. Geol. Soc. India*, 1929, **54**, 1–166.
2. Dunn, J. A., *Mem. Geol. Soc. India*, 1940, **63**, 303–369.
3. Jones, H. C., *Mem. Geol. Soc. India*, 1934, **63**, 167–302.
4. Krishnan, M. S., *Mem. Geol. Soc. India*, 1937, **71**, 1–181.
5. Dunn, J. A. and Dey, A. K., *Mem. Geol. Soc. India*, 1940, **69**, 281–450.
6. Sharma, M., Basu, A. R. and Ray, S. L., *Contrib. Mineral. Petrol.*, 1994, **117**, 45–55.
7. Bhattacharya, D. S., *Bull. Geol. Min. Met. Soc. India*, 1966, **36**, 1–21.
8. Bhattacharya, D. S., Pasayat, S. and Sarkar, A. N., *Indian J. Earth Sci.*, 1976, **3**, 26–36.
9. Chakraborty, K. R. and Sen, S. K., *Contrib. Mineral. Petrol.*, 1967, **16**, 210–232.
10. Mukhopadhyay, D., Bhattacharya, T., Chakraborty, T. and Dey, A. K., *Proc.*

Indian Acad. Sci. (*Earth Planet. Sci.*), 1990, **99**, 249–268.
11. Roy, A. B., *Geol. Mijnb.*, 1966, **45**, 365–374.

DHRUBA MUKHOPADHYAY

Department of Geology,
University of Calcutta,
35 Ballygunge Circular Road,
Calcutta 700 019, India

Annual Review of Phytopathology, 1994. R. James Cook (ed.), Annual Reviews Inc., 4139 El Camino Way, Palo Alto, California 94303-0897, USA. Vol. 32. 639 pp. Price: USA $49, elsewhere $54.

As a professional plant pathologist, the reviewer himself has been an avid and regular reader of this review series and has all the volumes in his own library. This series has been popular with plant pathologists all over the world because of its wide coverage of topics, ranging from historical to those covering the cutting edges of the science of plant pathology. Many topics chosen are of international interest. Volume 32 once again follows the rich tradition.

The volume has 30 review topics classified into 13 groups: Prefatory chapter; Pioneer leaders; Diagnosis and appraisal of plant disease; Pathogens: Fungi; Pathogens: Bacteria; Pathogens: Nematodes; Pathogens: Viruses; Physiology, morphology, anatomy, biochemistry and molecular biology of host–pathogen interactions; Epidemiology and influence of climate; Toxicants and chemical control; Biological and culture control; and Special topics.

The first two groups – Prefatory chapter and Pioneer leaders – relate to historical aspects. There is only one article under the first group written by Dr George A. Zentmeyer about his own 55-year innings as a professional pathologist. It is very readable and particularly useful to young plant pathologists, who could learn from the varied, exciting experiences narrated by the author. The second group has three articles on three pioneers in the field of plant pathology – Harry Marshall Ward, Tom Goodey and Frederick Charles Bawden – all from the United Kingdom, whose work has had tremendous impact internationally. There is an interesting account of how the rust disease of coffee wiped out the crop from Sri Lanka and was replaced by tea towards the end of the last century, despite efforts by experts like Harry Marshall Ward. Tom Goodey is today considered the father of nematology in Britain. Bawden's contribution to the field of virology in 1930s has been a landmark; he and his colleague Pirie provided for the first time evidence that viruses are nucleoproteins. All the three articles make very interesting reading.

The third group, Diagnosis and appraisal of plant disease, contains three articles. The first one deals with ash (*Fraxinus* spp.) yellows disease, the second with dogwood (*Cornus* spp.) anthracnose, and the third with the so-called beech (*Fagus* spp.) bark disease. The three articles will be of interest mostly to the scientists of North America and possibly Europe and other temperate regions.

The next four groups cover topics related to fungi, bacteria, nematodes and viruses as pathogens. There is a very interesting review on the presence of double-standard ribonucleic acids (dsRNA) in the rust fungi. The presence of dsRNA seems to be a normal feature of most rust fungi even though these fungi continue to be damaging pathogens of crop plants. In the next few years the role of dsRNA in the pathoginicity of rust fungi should become clear. Another article describes the molecular biology approach that might help in improving our taxonomic understanding of *Rhizoctonia*. There is an excellent review on our status of knowledge regarding the secretion of extracellular virulence factors by phytopathogenic bacteria. Another review worth reading is the one by Sijmons *et al.* highlighting the recent advances made in the areas of parasitic strategies and host cell responses. Again, modern molecular biology techniques have helped and will further help in improving our understanding of the host–root nematode interactions. Modern molecular biology techniques have contributed in rapidly improving our understanding of plant viruses. There is an interesting review on the evolution of closteroviruses, which are known for having by far the largest positive-strand RNA genomes among RNA viruses of plants. Another excellent review covers the topic of plant viral RNA synthesis in cell-free systems. Though our knowledge of initial stages of virus multiplication in plant hosts has increased substantially in the last two decades, it is far from being complete. Another review discusses our recently gained knowledge of RNA–RNA recombinations and evolution in virus-infected plants. Seed transmission of viruses concerns plant pathologists all over the world and this volume has an excellent review discussing the work mainly done in the last 25 years. Another review covers our knowledge of the role of plasmodesmata in viral transport in plants. Though annual reviews are normally edited very well, it is difficult to understand how the editors made the mistake of including the review on the molecular systemics and population biology of *Rhizoctonia*, a fungus, in the group of articles on bacteria.

In the remaining groups of reviews, the readers will find articles dealing with topics like (i) early events in the activation of plant defence response, (ii) social and political implications of managing plant diseases with decreased availability of fungicides in the USA and Europe, and (iii) the role of plant clinics in disease diagnosis and education in North America.

The volume includes a thoroughly prepared subject index. The reviewer is once again convinced about the outstanding service that the volumes of the *Annual Review of Phytopathology* provide to plant pathologists of the world.

Y. L. NENE

International Crops Research Institute for the Semi-Arid Tropics,
Patancheru 502 325, India

Professor Willy Fowler (1911–1995) – An obituary

William Alfred Fowler, Nobel Laureate (1983), passed away on 14 March 1995 in Pasadena, California, the place where he had worked for about six decades. He was better known to all his friends as 'Willy'.

Born on 9 August 1911 in Pittsburgh, Willy Fowler did his first degree at the Ohio State University and then in 1933 went to the California Institute of Technology as a graduate student of Charles Lauritsen at the newly established Kellogg Radiation Laboratory. With so much nuclear astrophysics that has come out of Kellogg, it is interesting to know that the lab had been established to study the physics of 1 MeV X-rays and their application to the treatment of cancer. As Willy recalled in his Nobel Lecture, the change of direction came right at the beginning with the discovery in 1932 by Walton and Cockroft that nuclei could be disintegrated by protons accelerated to energies below 1 MeV. This led Lauritsen to convert one of his X-ray tubes to a positive-ion accelerator and to initiate work in nuclear physics.

The shift towards astrophysics took place in Willy's postdoctoral career at Caltech, where, in fact, he remained throughout his life. His early work with his colleagues at Kellogg included the study of the proton–proton chain and the carbon–nitrogen cycle for making helium from fusion of hydrogen. Kellogg provided data on the many experimental cross-sections for the intermediate reactions in these processes. The fusion process from light to heavier nuclei, however, encounters a stumbling block beyond helium (atomic mass 4).

There is a mass gap from 5 to 8 showing a lack of stable nuclei in this mass range and this was the subject of Willy's study for several years. In the end it became clear that because of this gap the grand scheme proposed by George Gamow of synthesizing all nuclei in the early hot era of the big bang universe would not work. How then were heavier nuclei like carbon and oxygen, or iron, made? Astrophysicists building on the pioneering work of Eddington in the 1920s and Bethe in the late 1930s expected the answer to come from stars. But here also, how was the mass gap to be bridged?

The clue to this question came from a highly imaginative solution proposed by Fred Hoyle in 1953, which was to bring Hoyle and Fowler together in a very fruitful collaboration in nuclear astrophysics, a collaboration that lasted nearly 30 years. It was Hoyle's conjecture that the element-building process would more naturally take place inside stars whose cores were known to be hot enough to sustain nuclear fusion reactions. And to bridge the mass gap after helium, Hoyle proposed that the process would jump over the unstable gap. Thus, three nuclei of helium (atomic mass 4) would fuse to make carbon (atomic mass 12) in a resonant reaction. Resonance would make the triple fusion go fast, to compensate for the rarity of three nuclei coming together.



Willy used to recall how Fred walked into Kellogg with this idea which led to the verifiable result that there ought to exist an excited level of the carbon nucleus to which the three helium nuclei get converted. Look for it, he urged the nuclear physicists. After an initial scepticism, Ward Whaling and others at Kellogg got going and did indeed succeed in confirming Hoyle's conjecture. As the helium nucleus is called the alpha particle, this process came to be known as the *triple alpha process*.

This episode led Willy firmly and irrevocably into astrophysics, where plenty of challenges existed for nuclear physicists, especially in the area of stellar evolution. Eddington's work had shown the way of constructing stellar models essentially for the main sequence stars, whose energy is generated from the fusion of hydrogen to helium. What next? Hoyle's triple alpha process had opened the way for the synthesis of even heavier nuclei. But information on the various reaction rates was lacking. Willy turned his attention to this problem.

While Hoyle provided the theory on stellar structure, Fowler brought the nuclear physics inputs. Still lacking were the theoretical and observational details on how radiation interacts with matter as it escapes from the stars. The husband and wife team of Geoffrey and Margaret Burbidge provided this help in a fourfold collaboration that came to be known as the B^2FH, which culminated in 1957 in a mammoth paper entitled 'The synthesis of elements in stars' in the *Reviews of Modern Physics* (**29**, p. 547). This B^2FH was to set the tone for all future work in stellar nucleosynthesis, for it covered the synthesis of different nuclei in stars in all their different stages of evolution.

From the late fifties to the mid-sixties the Hoyle–Fowler combination turned from stellar setting to the larger extragalactic one. Radio astronomy was beginning to provide glimpses of the highly energetic extragalactic world. Geoffrey Burbidge had been the first one to draw attention to the fact that taken at face value the powerful radio sources demanded huge energy reservoirs that far exceeded the dynamical energies of colliding galaxies. Some new source was needed. Would the stellar option, say, of a supermassive star work? Fowler and Hoyle considered the models of such objects and came to the conclusion that these objects will have a far more dominating force to contend with, namely, the force of gravity. Under such a force these objects would tend to shrink faster and faster. In a paper in *Nature* in early 1963 (**197**, p. 533) they conjectured such highly collapsed supermassive objects as likely sources of the observed energy.

Parallelly, the observational front was also progressing and, thanks to a rare collaboration between the optical and radio astronomers in 1962–63, the first discovery of the quasistellar objects (QSOs) was made. Almost star-like in appearance the QSOs showed large redshifts in their spectra and theoretical estimates led to the conclusion that

energy comparable to that of an entire galaxy might well be coming out of a source no greater than a light year across. Here then was the very example of the kind of objects conjectured by Hoyle and Fowler.

This discovery opened up a new field in astrophysics, called *relativistic astrophysics*, which dealt with the astrophysical phenomena in strong gravitational fields, in situations where Newtonian gravity must be replaced by Einstein's general theory of relativity. So great was the excitement amongst the astronomical community that an international symposium was held in December 1963 in Dallas to discuss the implications of theories and observations in this emerging field. Both Hoyle and Fowler were the lead speakers in this meeting on the theoretical side. These meetings now continue to be held biennially under the title 'Texas Symposia'. Much of the present black hole bandwagon in astrophysics has its origin in the Hoyle–Fowler work of 1963.

Another of Willy's seminal contributions came in the big bang cosmology when in 1967 he along with Fred Hoyle and Robert V. Wagoner carried out a revised and updated version of Gamow's primordial nucleosynthesis. The Wagoner–Fowler–Hoyle paper in the *Astrophysical Journal* (**148**, p. 3) again turned out to be a trendsetter for future work in big bang nucleosynthesis. It is interesting to recall that this work was carried out in a shed in the Cambridge Observatories while Hoyle's new institute building was under construction.

It was during the first six years of the Institute of Theoretical Astronomy at Cambridge that Willy was a frequent visitor during the summer months. He and Fred would sometimes take off to the Scottish Highlands for hiking, a practice that Willy continued till late in his life.

For his work in nuclear astrophysics Willy Fowler shared the 1983 Nobel Prize with S. Chandrasekar. It was the second time that a Nobel award was given for theoretical astrophysics, the previous occasion being in 1967 when Hans Bethe was honoured for his work on stellar structure models with nuclear energy generation. The 1983 award, however, brought both surprise and disappointment at the omission of Fred Hoyle from the list.

Willy was known for his thoroughness in all the work he did, a fact which was belied by his informality and jolly demeanour. He would be the life and soul of a party discussing (over a martini) either the world series, or the Scottish terrain, or one of his long train journeys (he was a train buff), or nuclear cross-sections all with equal ease. Amongst his many honours, his Indian connection was with IUCAA as one of its Honorary Fellows, and several of us recall his jokes and anecdotes as he talked about his latest work in cosmology in March 1990. As a personality he will be greatly missed but his work will continue to guide the succeeding generations of astrophysicists.

Willy Fowler is survived by his wife Mary Dutcher Fowler and two daughters by his first wife Ardienne (deceased).

JAYANT NARLIKAR

Inter-University Centre for Astronomy and Astrophysics,
Post Bag 4, Ganeshkhind,
Pune 411 007, India

GENERALIA

Who should look at stars*

Ashok S. Ganguly

In January 1994, I met Jayant Narlikar on a flight from Calcutta to Bombay. He then reminded me of an earlier occasion when (in a weak moment), I had agreed to deliver a Foundation Day lecture at his Institute in Pune. I tried to wriggle myself out of the commitment, firstly, by reminding Jayant that I would have nothing meaningful to say to what was likely to be a distinguished gathering of scholars, and secondly, because my daughters were discouraging me from being typecast as a Foundation Day fixture. Both attempts failed, as you will notice from my presence here today.

Having made the commitment to speak, I spent many agonizing evenings trying to focus on a suitable topic. Finally, I drew inspiration from my association with Narlikar as a member of the Science Advisory Council to the Prime Minister (1986–89) and settled on a 'blue sky' approach. More of that shortly.

A few recent events spurred me to put down some initial thoughts on paper before I lost track of them. A very close friend and his wife were visiting us in London. While we were driving to Glyndebourne to attend a performance of the Opera, Don Giovani, I shared with them my dilemma about the Foundation Day lecture. My friend's wife, a well-known physician from Bombay in her own right, asked me, 'Why should a poor country like India, with its impoverished millions, fund research to watch stars?' For the best part of the two-hour car ride, we had a series of disagreements on the subject. These disagreements ranged from the concept of curiosity, through the nature of the Indian mind and the need to be part of the international community in leading-edge science, to the origins of the universe and life on Earth. I could not convince her that curiosity-driven exploration was at the heart of human civilization, starving or otherwise. Finally, the arguments ended inconclusively when I raised a rhetorical question, 'What was the level of poverty and deprivation in Italy in the days of Galileo Galilei?', I asked, feeling that would end the debate amicably. Neither

*This was the subject of the Foundation Day Lecture delivered at Inter-University Centre for Astronomy and Astrophysics, Pune, India, on 29 December 1994.

of us knew for sure, but she promptly shot back that it really did not matter, since Galileo's star gazing was not funded by the State.

Galileo – the star gazer

Her reply had me stumped, until just a few days later when I came across a review by A. C. Grayling of GALILEO: A LIFE, a new book by James Reston Jr. If any one man acted as a doorman to the modern world, Galileo Galilei has an excellent claim to the title. The publication of his latest biography is timely because Reston has hooked his tale firmly to events in the contemporary world: the flight of the spacecraft Galileo to Jupiter, still taking place as I speak to you, and the decision of the Vatican in 1992 to acknowledge its fault in ill-treating Galileo three-and-a half centuries ago by dragging him before the Inquisition, humiliating and imprisoning him.

Indeed, Galileo's story is a microcosm of the epic struggle between science, religion and society. Galileo was inquisitive, inventive and mathematically adept. He was fascinated by the view above him in the clear Italian night sky. He perfected the telescope, and terrified the Church by revealing more stars than had been guessed before, hitherto unseen satellites orbiting other planets, valleys and mountains on the surface of the moon.

Because Christian scriptures taught that the Earth sits immovably in the centre of the Universe, whose celestial spheres are driven around by angels, the Church could not tolerate this new cosmology.

By threats and intimidation, Pope Urban VIII forced Galileo to recant his espousal of the Copernican system. That the Church should come to its own recantation in 1992 speaks volumes for the conflict between faith and reason. Therefore, I cannot consider my friend's wife's question of whether astronomy research is really needed in a poor country like India unusual.

Galileo not only made discoveries of prime importance in astronomy and physics – especially in the laws of motion, thus breaking the stranglehold of Aristotelian ideas – he was also an inventive genius. He devised, amongst other things, pendulums for clocks, ways of improving telescopes and instruments for measuring pulse rates and temperatures. Despite disgrace by the Inquisition, his telescopic discoveries made him a star all over Europe.

While I have not been able to locate evidence of State support for his experimental pursuits. I am reliably informed that there were many poor people in Italy at the time Galileo gazed at the stars.

From star-gazing Galileo to the man behind the comet, Edmund Halley

Soon after the book on Galileo, I came across an excellent piece by Allan Chapman of Wadham College, Oxford, on a remarkable man.

Chapman avers that scientific research has always been expensive, but in the past, the real cost was not so much the hardware, as paying for the scientists' time. This is why so many discoveries between the 17th and the 19th centuries have been made by clergymen, semi-retired professionals and businessmen, and people of private means – individuals who were in command of their own time. But while the cash was not entirely certain, if one had the ability to wear many hats and cultivate friends across the whole spectrum of potential influence, then one might enjoy a career similar to that of Edmund Halley (1656–1742). This immediately brought to my mind some eminent Indians who fit parts of this description, like Ramanujam, C. V. Raman and Homi Bhabha, to name just three.

Edmund Halley's work was funded by, to begin with, family money. When that ran out, he worked as a clerk at the Royal Society, then took a job with the Royal mint, and later joined the Royal Navy. At the age of 48, he became a professor at Oxford and at 64 the Astronomer Royal. Edmund Halley was clearly capable of living with uncertainty.

Though primarily an astronomer, what is most impressive about Halley's approach to science is his awareness of how the great 'systems' of nature are interconnected. Geomagnetism was such a system. Halley published major papers in 1683 and 1692 in which he collected and analysed a growing body of data on the subject. He created the geomagnetic chart and then in 1716 he was the first to recognize the interconnection between the Earth's magnetic field and the aurora borealis.

Halley was a man of vast interests, ranging from pure science to a whole series of practical applications in marine engineering.

I imagine in modern parlance one would find comparable traits in the venture entrepreneurship of many famous American academics and Nobel Laureates. But where are the young Bhabhas and Ramans, I wonder.

Fred Hoyle – the astrobiologist

No address in an institute headed by Jayant Narlikar would be complete without reference to his Guru, Fred Hoyle, the astrobiologist.

Being somewhat of an amateur biologist myself, Hoyle's 'origin of life' hypothesis intrigues me. In his autobiography, *Home Is Where The Wind Blows*, Hoyle is sanguine about the chances of acceptance of his idea that life began in the vast expanses of space. 'Finding glycine is about 5% of the way to proving the idea', says Hoyle, 'It's the right way, but it's only 5%'.

The glycine discovery certainly bolsters Hoyle's thesis that the first organic matter rained down on Earth from space. According to Hoyle, it was such a comet which struck Earth about four billion years ago, depositing its cargo of primitive cells – the forerunners of all life today.

It is a rather controversial view, although the idea that organic material – the 'feedstock of life' – came from space is gaining wider acceptance. One is, however, faced with the moral and philosophical dilemma about the true origin of the living form. Hoyle acknowledges that the riddle will not be solved in his lifetime. However, he has been proved right before. In 1940 he had suggested that molecular hydrogen was widespread in space and was greeted with wide disbelief. He subsequently published his views in a science fiction book, *The Black Cloud*. Since the publication of the book, more than 100 molecules have been detected in space, including molecular hydrogen, which is now known to be the most abundant molecule in the Universe.

Marcus Chown (*New Scientist*, 10 September 1994) writes, 'Hoyle believes that, if a problem has an orthodox solution, the scientific community would already have found it, so he looks for the unorthodox solution'.

This unorthodox approach has proved Hoyle right on a number of predictions and wrong in just as many, such as the making of the sun or the explanation of binary stars.

But as the saying goes, 'If you have not made mistakes, you have really not tried'. I hope the adage applies to this Institute also.

I was pleased to learn that since 1988 Hoyle has developed the 'quasi-steady-state theory' jointly with Burbridge and Jayant Narlikar. The three of them believe that the creation tap opened in one part of the Universe 15 billion years ago, unleashing a flood of matter and causing the expansion of galaxies we observe all about us. Who knows they may be right!

Finally, I feel much comforted by Hoyle's view that 'a little bit of god operates in all of us. We are his observing instruments. He observes the Universe through us'. No matter where you are, whether in Cambridge or California or Pune, it all happened together.

Has science lost its way?

While I was warming up to the theme of science for curiosity's sake, I was crestfallen to read Bernard Levin's essay titled 'Has science lost its way?' (Even if I were to ignore Levin's levity, in the UK, observations on science such as his fall under the classification of 'Have you beaten your wife today?'.) Like it or not, some of this reflects contemporary lay thinking.

Bernard Levin's motto seems to be 'Beware of scientists claiming breakthroughs'. He then goes on to observe that 'If the boffins do have a fault, it is their conviction that any amount of money required for their work should be immediately provided'. In this essay, Levin chose the quarks as his centrepiece to criticize the nature of scientific enquiry and quotes a neat example of his odd but touching belief from a recent breakthrough in science. The last quark had been spotted (or rather not spotted. The thing is invisible). Arie Bodek from Chicago said that his discovery had saved scientists from an intellectual crisis 'since failure to discover the top quark would have shattered decades of research worth billions of pounds'. Levin asks, if the quark cannot even be seen, 'What is in it for us?'. The moot question is, what can discovery of quarks do for the common man? To Levin's queries, the only response he was able to elicit is that there are six quarks named UP, DOWN, CHARM, STRANGE, TOP and BOTTOM, and that 493 scientists were enlisted to find the last quark, going about the search 'smashing protons and antiprotons in a four-mile circular accelerator' and on top of that 'the two particles annihilate each other as they collide'.

Bernard Levin then moves on to compound his confusion by relating the discovery of a new planet 7000 trillion miles away. Hitherto 'in every case, the putative planets had either proved impossible to confirm, turned out to be something else, or shown to be a product of error'. Now, however, Wolszczan's team 'has bolstered its original findings with three more years of even more finely nuanced data designed to eliminate any explanation other than planets'. And this statement is backed by Kulkarni, a pulsar expert at the California Institute of Technology, who says, 'It should convince even die-hard sceptics that planets exist outside the solar system'.

Amongst other things, Levin concludes by noting that, whereas the common man may not very much begrudge the scientists the millions and millions of pounds needed to peep beyond the solar system 'where God resides!' while millions and millions may starve and suffer from malnourishment around the world, surely the tax-payers have a right to understand better where all these discoveries are leading.

Well, there we go! In quick succession, the wife of one of my best friends from India and the well-known columnist Bernard Levin, two persons from two totally different societies, asking very similar questions. Then interesting events started happening in July.

Shoemaker Levy 9

My interest perked up on reading the news that the comet Shoemaker Levy 9 was moving as a funeral procession of 21 fragments of ice and rock after its break up two years ago. Their eventual collision with Jupiter was being described by astronomers as a 'once in a millennium opportunity'. The subsequent pictures and descriptions of the collision were widely reported around the world by the media. I am told that it will take years before the explanation of the events leading to this collision are well understood. But the pictures and descriptions of the events started making sense to the common man as to why we stare at the night sky. This has been further encouraged by the fact that, unlike the Shoemaker couple who are professional astronomers, Levy is an amateur.

Suddenly Big Bang is back on the agenda and the question on every one's lips is, 'If it could happen to Jupiter, can it not happen to Earth?'.

Comets are celestial bodies, believed to be the relics of the birth of the solar system. A thousand or so of them have had their orbits well plotted. But for every comet which is confirmed to exist, there are thought to be at least ten others that have eluded detection. And billions more may reside in a halo beyond Pluto. Comets have collided with Earth before, and scientists insist that more collisions are probable, making it vital to overcome complacency about the threat to Earth from cosmic debris.

If a comet, streaming along at 50,000 miles per hour, were to crash into Earth in the wrong place at the wrong time – London, Bombay or New York at rush hour – it could cause a disaster of monstrous proportions. These thoughts were expressed in the editorials of the serious *London Times*. There is a school of serious astronomers, aptly known as 'catastrophists', which is convinced that within the celestial procession of comets lie the seeds of Armageddon. They attribute the death of the dinosaurs, 65 million years ago, to a comet which ploughed into the planet triggering either a nuclear-style winter or uncontrollable fires. The serious question is: If the comet could wipe out the mighty *Tyrannosaurus rex*, what chance do puny *Homo sapiens* have?

Thus, contemplating the impact of the 21 fragments of Shoemaker Levy 9 on collision with Jupiter in our lifetime has raised modern astronomy from the preoccupation of the specialists to a concern for all living and thinking mankind, starving or not.

Such an extraordinary astral event then starts stretching the mind and arouses the interest of the layman in astronomy, as indeed it did mine. I then began to think that although man has the technology needed to detect and track threatening celestial objects, there is yet no global chain of telescopes. And research has still to be aimed at defending the Earth from 'asteroid attack', including the possible redeploy-

ment of nuclear warheads to shatter incoming projectiles and the use of propulsion units to divert or steer them away. None of these, however, will protect us from every conceivable threat, nor can they whet our awe at the immensity of space. The 'black that remains beyond our blue' seems destined to remain a mystery to all but our telescopes.

The social relevance

No matter how awestruck I am, I cannot avoid relating all of this to developments in India. I hope I have provided a reasonable, layman's support to the need for the pursuit of the planets even in Pune. The statistical probability is that, next to China, India should have the largest number of curious people. This does not necessarily mean that we should have the second largest population of astronomers peeping through a chain of telescopes across the Indian subcontinent. However, in astronomy we have a history of interest as recorded in the Puranas, through the Middle Ages, and now this world-class centre in Pune. There is, thus, a tradition in mathematics and physics which links the ancient with the leading edge in India.

This, unfortunately, is not the case in most other areas of modern science and technology. Although we talk of modernizing and globalizing India, in lay terms what we really mean is a growth in production and consumption in our lifetime rather than a change in the fundamentals of looking and doing for a better future. I am, therefore, proud that India is considered to be a world player in astronomy and astrophysics, if not in any other sphere of science and technology.

Following our independence after over a thousand years of enslavement, Nehru, Bhabha, Bhatnagar, Meghnad Saha and many others inspired my generation to believe that science and technology held the key to traverse the path from backwardness to modernity and provide social justice for all. History will record that for the first three decades this promise at least seemed to begin to materialize. The role of science and technology in India achieving self-reliance was indeed significant. India thus achieved a critical mass as a democracy with a viable economy, although the problems of social inequity remained.

At a time when we are ready to integrate India's economic developments with the rest of the world, the same science and technology institutions which helped in achieving self-reliance may now turn out to be the major source of competitive disadvantage. Any attempt to revive them may be futile, as many of them may have outlived their utility. What is necessary is to take drastic steps to restructure and modernize the old and, in addition, create new institutions for the future.

Who should be concerned about this vital problem? The Government, Indian Industry, the non-Government Organization (NGOs), the media, the intellectuals, or whoever? Since 1990 even the interest and vibrancy in the science and technology community that one felt and saw in the 1980s seems to have dangerously waned. Unusual and dangerous priorities have overtaken the nation, such as disputes regarding places of worship, the revival of the caste conflict, the drive for instant gratification and enrichment, multinational baiting, the 'back to primitive roots movement', etc. There is not even lip service paid to the social and economic imperative to modernize our institutes of higher education and those institutions dedicated to advanced research in different areas of S&T. Every country other than India seems to be deeply engaged in addressing these vital issues. China is set to revamp its total research base, from one which had remained isolated and totally dependent on the State, towards one which is more autonomous, modern and globally relevant. There is a similar reassessment of the science base in the USA and Japan. In the UK, the 1993 White Paper on the need to choose areas of science in which to lead and its relevance to that country's global competitiveness was articulated. The key was that not every country could afford to be outstanding in every discipline of science. What a nation needs is a solid foundation in basic and higher education and then the ability to choose areas of science to excel in, by the combined efforts of the Government, Academia and the Industry.

As an Indian it frightens me today that, as we enter the most complex and competitive period in human history, we in India seem to adopt a policy of benign neglect of higher education and leading-edge S&T. Probably, the only other tragedy that this policy can be compared with is the disarray in the once powerful S&T culture of the former Soviet Union with its devastating consequences.

India does not have a choice. All of us must raise our voices against this benign neglect, in which lie the seeds of economic impoverishment and social disarray. Just talking about liberalization, economic development and poverty alleviation will not do. There are limits to foreign investment and export growth. Without overhauling our S&T base India faces a finite economic horizon.

I, therefore, disagree with my friend's wife and with people like Bernard Levin. Just because we are unable to comprehend science-speak or because many mediocre practitioners may have given science a poor image, it does not mean that there are better alternatives to achieving excellence in science. For there are none. More people will die of starvation and malnourishment without scientific advances and more nations will face decline and catastrophe without the pursuit of scientific knowledge. I am, therefore, thrilled that Narlikar and his dedicated colleagues have built this world-class facility in Pune. Could this be the faint glimmer of light at the end of the tunnel? We need a few more of these, and soon. I am encouraged by developments in massive parallel computing, also in Pune, advances in complex materials in the Institute of Science in Bangalore, the modernization of NCL and a few other such developments. They cost very little money. What they need is a few great minds. As an Indian, I am grateful that we have a fair number of excellent minds. Let us create the environment in which they can flourish to prepare India for the next century.

Jayant and I had the privilege of being members of Rajiv Gandhi's Science Advisory Council. In every meeting we had with him, we enjoyed inspirational debates on the choice of blue-sky priorities. True, even then, everyone was aware of India's day-to-day problems, but it was not a question of either/or. Such debates have now waned. I hope Narlikar and his dedicated group will help revive the debates, even beyond blue skies. I sometimes wonder, if only we could divert a fraction of the energy of our people away from breaking down places of worship or participating in caste riots, towards the pursuit of scientific enquiry, could we transform our nation and start realizing our potential?

I am now pleased about the brief encounter Jayant and I had in January 1994.

Ashok S. Ganguly works in the Research & Engineering Division, Unilever PLC, London, UK.

Scientific tradition and other traditions*

B. M. Udgaonkar

In the 'Project of History of Indian Science, Philosophy and Culture' (PHISPC), our concern has to be with the trio science–philosophy–culture, in interaction with one another. This is where the uniqueness of this project would lie. Unidimensional discussions of any one of the trio should have no place in the volumes that result, except where they raise important questions at the interface(s). And in that case, experts in two (or three) dimensions should have sessions together to sharpen the questions and, to the extent possible, arrive at (even tentative) answers to these. This has already been emphasized by Chattopadhyaya[1]. However, it bears repetition in view of the contents and presentation in some of the 'Occasional Papers' that have been brought out so far.

Indian civilization is about 4500–5000 years old (including the Indus Valley Civilization period; 8000 years old if one includes the pre-Harappan findings in Baluchistan). During this period, India has made many important contributions to science, technology, philosophy and culture. How did the developments in each of these areas influence one another? During this long period, what was the place of science in philosophy and in the civilization at large, at different epochs?

What role do the dominant scientific concepts play in the differential development of philosophy in different cultures? For example, Seal[2] observes that 'Hindu philosophy on its empirical side was dominated by concepts derived from physiology and philology, just as Greek philosophy was similarly dominated by geometrical concepts and methods', and adds that 'Comparative philosophy, then, in its criticism and estimate of the Hindu thought, must take note of the empirical basis on which the speculative superstructure was based'. This assertion would merit a detailed discussion. Is the assertion true? Or is it an oversimplification with regard to both Hindu Philosophy and Greek Philosophy?

Established science at a given time influences philosophy and *vice versa*, though radically new developments do so usually with some time lag. How does one see these influences at work, in general in the Indian context? What were the positive and negative influences at different epochs? Chronology becomes important in this context. There are, however, serious chronological uncertainties in Indian history.

In the case of science, one tends to compare the Indian situation with the developments in Europe over the last few hundred years, is overawed, and it is not uncommon to hear loaded counterfactual questions like 'Why did India have no science? Is it because of its tradition?'. In asking such questions, the foreigner and, following him, many Indians, forget that science based on observation and empirical methods is to be found in the practice of the ancients in India, and that science is as much a part of our heritage as the greatly extolled mystical tradition[3–5]. One has legitimately to take a longer time scale for any comparison, and counter the ethnocentric prejudices with facts.

The time scale of several thousand years also means that concepts may have changed with time. This has to be constantly kept in view. For the impact of any given concept, e.g. *Māyā*, one has to ask the meaning attached to it during different epochs.

Indian contributions to science and technology from the earliest times have begun to be documented[3–5], though there is no comprehensive work yet comparable to that of Needham and coworkers on China. It is, however, by now recognized that India's contributions to science and technology over the centuries are quite significant. This is, however, not the subject of our discussion. Rather, we will try to survey the reasons which have often been adduced as to why, in spite of a strong early base in science and technology, India failed to build on it during the last 500 years. Why did this happen?

Why did India lag behind in recent centuries?

The question often raised is: If Indians had the questioning attitude and the tradition of experimental and mathematical investigations at one time, and if they made significant contributions to developments in science and technology, especially up to about 1200 AD, and maintained their superiority in certain technologies, e.g. textiles, metallurgy, architecture, ship building, etc., right up to the 17th or 18th century[6–8], why did they fail to build further on this base? Why was it left to Europe to develop modern science? Were there some factors in our traditions which came in the way?

In history, such questions as to why something did not happen are difficult to answer. They are even treated as nonquestions by some historians. There are several questions of this kind that one could ask about Europe: e.g. why is it that none of the religions of today started in Europe? Why did they all start in Asia? Why did the early wave of science and philosophy in Europe not start in mainland Europe? Why did it start in an insignificant southeastern corner of Europe and in adjoining Asia (called Asia *minor*)? Why did the Europeans not discover zero? And why was Japan *relatively late* in joining the march of civilizations? And so on. One has no simple and straightforward answers to such questions.

China and the Arab world were also unable to build on the early head-start. Maybe different sociopolitical and historical circumstances were responsible in each case. These need to be studied in an unprejudiced manner, using such meagre data as are available and reconstructing where there are gaps (and there *are* many gaps), with a more open mind than has been customary. Many times hypotheses were made by Western scholars in the colonial period, and these have been echoed by Indian scholars even in the post-Independence era. A kind of colonial thinking has tended to persist, even if unconsciously, and so these hypotheses have to be looked at afresh.

*This article was originally prepared for the Project of History of Indian Science Philosophy and Culture of the Indian Council of Philosophical Research and was brought out as Occasional Paper 24 of the Project.

While listing some of the ideological or attitudinal factors which have been held responsible for India's inability to build on its early base, and making some comments on them, we would like to emphasize at the start that they do not strike us as having the status of anything more than speculations or prejudices, given the importance of the question being discussed. Much more detailed work is needed before any of the prevalent myths in this domain could attain the status of causal relationships *vis-à-vis* development of science in India. Wherever any of the factors have their counterparts in other civilizations, especially the European civilization, a comparative study would be called for.

It is to be hoped that the relevant sections on philosophy and culture in the volumes that will come out of PHISPC will deal with these questions in depth.

The supposed other-worldliness of Indian culture

It is alleged that early Indian culture was other-worldly, that it perceived the world as *Māyā* or illusion, and that the conceptual world-worthlessness led to a lack of interest among Indians in any worldly pursuits. Indians, it is said, were only interested in the liberation of the soul as a goal, and not in studying the external world.

It may be argued that a distortion of the concept of *Māyā* is involved in such an argument. Even if this is so, it would be useful to know the extent to which the distorted/corrupted interpretation of *Māyā* affected a young intellectual's choice of a field of study during different epochs of our long history. Were the brightest enquiring minds turned away from the study of worldly phenomena?

Related to the interest in other-worldly entities was the distinction made between what was called *parā vidyā* and *aparā vidyā*, the higher kind of knowledge and the lower kind of knowledge, and the influence such distinction would have had on the perception of what was really worth studying and worth respecting.

In the *Mundaka Upanisad*, we find the question put by an earnest student to a great teacher: 'What is that reality by knowing which we can know all that there is in this manifested universe?'. The teacher replies: 'Two are the *vidyās* or sciences to be acquired by man, so say the knowers of *Brahman*. One is called *parā vidyā*, higher science or knowledge', the other is called *aparā vidyā*, ordinary science or knowledge'[9]. The knowledge/search for the knowledge of the soul was considered the higher knowledge. Since what was considered *really worth knowing* was to be approached by turning the light inwards, this may have come in the way of the study of the external world. The best minds may have turned towards the exploration of the inner world – toward spiritualism and mysticism – which was placed on a higher pedestal in the name of *parā vidyā*.

Then there was 'a powerful current'[10] in Indian culture that 'the truths of the physical and supra-physical realities can be grasped, known and possessed by us through faculties which lie above the ranges of physical senses and rational intelligence'[10] and that the knowledge obtained by development of those higher faculties, by yogic means, through meditation and intuition, is superior to that obtained by one's senses and reason[11]. The gods have sometimes been described as *pratyaksa-dvisah*, i.e. as hating something that is acquired by the senses. How powerful was this current *in space and time*?

It has also been stated that the fact that renunciation was an accepted and respectable way of life fed these trends and came in the way of development of science and technology.

The *Karma* ideology has also been adduced as a cause – the idea of previous births and their consequences in the current life, and the resultant fatalism.

The influence of the *Bhakti-Marga*, which spread in the country through the emergence of a series of saint-poets after the eighth century in different emergent languages and their compositions, which became very popular, has also been mentioned.

While Indian culture has had such negative tendencies, it is not clear how widespread was their effect in space and time, in a large country and a stratified society whose intellectual history spanned three to four thousand years – especially as India has many important scientific discoveries to its credit over the centuries. This is a matter that needs to be looked into carefully from the point of view of possible impact on the development of science. Vivekananda, one of the greatest of the modern exponents of *Yoga* and the *Vedānta*, has, for example, laid repeated stress on the experimental character of *Yoga* and on basing it on reason[12].

Referring to the flourishing trade between India and Rome during the period 200 BC to 300 AD, Thapar[13] mentions Pliny's complaint that the Indian trade was a serious drain on the national income of Rome. The Frenchman Bernier, who travelled in India more than 1500 years later (1656–1668), says: 'It should not escape notice that gold and silver, after circulating in any other quarter of the globe, came at last to be absorbed in Hindustan'[14]. These are not observations supporting the other-worldliness of Indians over the centuries.

It is interesting to see what Vivekananda has to say about *Māyā*, especially in relation to scientific pursuits. Vivekananda refers to *Māyā* as 'one of the pillars upon which the *Vēdānta* rests'[15]. He, however, warns us that the word *Māyā* is used incorrectly to denote 'illusion or delusion or some such thing'. Having said this, he proceeds to give a number of examples to illustrate what *Māyā* is, a couple of which are quite intriguing:

> 'A man of science rises, he is thirsting after knowledge. No sacrifice is too great, no struggle too hopeless for him. He moves onward discovering secret after secret of nature, . . . and what for? . . . Why should we give him glory? Why should he acquire fame? Does not nature do infinitely more than any human being can do? . . . Nature can hurl a thunderbolt of any magnitude to any distance. If a man can do one small part as much, we praise him and send him to the skies. Why? Why should we praise him for imitating nature, imitating death, imitating dullness, imitating insentience? . . . Yet we are all struggling after that. And this is *Māyā*'[16]. '. . . so with our intellect. In our desire to solve the mysteries of the universe, we cannot stop our questioning, we feel we must know and cannot believe that no knowledge is to be gained. A few steps, and there arises the wall of beginningless and endless time which we cannot surmount. A few steps, and there appears a wall of boundless space which cannot be surmounted, and the whole is irrevocably bound in by the walls of cause and effect. We cannot go beyond them. Yet we struggle and still have to struggle. And this is *Māyā*'[17].

'. . . Is there no hope then?' He continues, 'True it is that we are all slaves of *Māyā*, and live in *Māyā*. Is there no way out? No hope? . . . on the other side, therefore, is the bold assertion that this is all nonsense, that this is *Māyā* . . .'[18].

At another place, Vivekananda says,

'. . . the minds of the people from whom the *Vedas* came were intent upon following principles, discovering principles. They had no time to work upon details or to wait for them, they wanted to go deep into the heart of things. Something beyond was calling them, as it were, and they could not wait. Scattered through the *Upanisads*, we find that the details of subjects which we now call modern sciences are often very erroneous, but at the same time their principles are correct . . .'[19].

We have chosen to quote at length from Vivekananda since he represents in a sense a link between ancient tradition (rather, the *Advaita* tradition) and modernity, and an important figure in the Indian renaissance towards the end of the last century. Is he also undervaluing the importance of scientific pursuits as we understand them? Do such views reflect the thinking about science in one important school of thought over the centuries? If so, could they have been partly responsible for nondevelopment of modern science in India? In science, what the Swamiji called 'details' are as important as 'principles'; 'details' are, in fact, an important path to 'principles' and have often led to the overthrow of earlier 'principles'. What does one mean by saying that details were erroneous while principles were correct? In science, no general principle is frozen as unassailable knowledge for all time. The strength of the scientific method is precisely in the openness to new principles in the light of new 'details' as they unfold themselves.

Thapar[20] has noted that 'Indians in the pre-eighteenth century never claimed that they were more spiritual than other people, or that the Indian way of life was concerned solely with things spiritual. The ideal of the earlier texts that the best life on earth is a balance between *dharma*, *artha* and *kāma* shows a healthy unconcern for any obsession with either the spiritual or the material'[20].

Works like the *Kautiliya Arthasāstra* show the detailed interest taken by ancient Indian scholars in worldly matters of government (which was not dominated by priesthood) in all its aspects.

Nehru has remarked: 'Probably the Semitic culture, as exemplified in many religions that emerged from it, and certainly Christianity, was far more otherworldly'. 'Indian culture taken as a whole', he adds, 'never emphasized the negation of life, though some of its philosophies did so; it seems to have done so much less than Christianity'[21]. Nehru points out that 'in India, we find during every period when her civilization bloomed, an intense joy of life and nature, a pleasure in the act of living, the development of art and music and literature and song and dancing and painting and the theatre, and even a highly sophisticated inquiry into the sex relation. It is inconceivable that a culture or view of life based on otherworldliness or world-worthlessness could have produced all these manifestations of vigorous and varied life. Indeed, it should be obvious that any culture that was basically other-worldly could not have carried on for thousands of years. The confusion seems to have arisen from the fact that Indian thought was always laying stress on the ultimate purpose of life. It could never forget the transcendental element in its make-up . . . It taught detachment in life and action, not abstention from them'[21].

Thapar[22] has discussed the origin of the stereotype. 'The theory of the instinctive unconcern of Indians for worldly things is found particularly useful in minimizing such obvious factors as colonialism in economic underdevelopment. The explanation of Indian economic backwardness does not lie in simplistic generalizations on Hindu culture but demands a detailed investigation into the nature of pre-colonial and colonial society in India'[22].

Alvarez[23] has observed: 'There is still no evidence to date that one line of thought was so dominant that it had become "orthodox": there was no single established view that set itself up against and above a number of other supposed heresies. The hypothesis that Hindu philosophy (wrongly identified wholly with the idealistic school) proved disastrous for India's technical capacity or economic activity should be relegated to the garbage heap; so too, the fashionable cliché that India spent too much of its energy in the pursuit of metaphysical speculations . . . .'

Suppression of the scientific spirit for political reasons

Chattopadhyaya[24] has argued that there was a systematic suppression of materialists – e.g. the *Cārvāks* – and of materialistic ideologies in ancient India. He has written extensively on this theme, and he, and before him Fillozat, has made a case study of medicine in ancient India.

Chattopadhyaya points out that already in the ancient period, some 2500 years ago, Indian medicine had taken the momentous leap from magico-religious therapeutics to rational therapeutics (from *daiva-vyapāsraya-bhesaja* to *yukti-vyapāsraya-bhesaja*). This was a very risky step, he says, the risk involved being political. 'To question even by implication – mysticism, ritualism and religion – necessitates also rejection of the very way of life all these tend to justify. It amounts to questioning the very norm on which a hierarchical society wants to thrive.' Further, the early medical men were looking for material, natural causes of diseases. Chattopadhyaya argues[25] that the view of the laws inherent in nature – technically called *svabhāva* – is essential for *Ayurveda*. Thus, the fundamental postulates of ancient Indian medicine are:

(i) everything in nature takes place according to laws inherent in nature;
(ii) though immutable, these laws are nevertheless knowable;
(iii) the knowledge of these laws brings power over nature, which medically means ensuring a long and healthy life.

This emphasis on *svabhāva* rejects *adrstavāda* or *karmavāda* or any supernaturalism. Chattopadhyaya[26] argues that this natural cause and effect approach to human ailments and their treatment was considered subversive by the priests, and the Indian law-givers were 'obliged to denounce it as abject heresy'. This led to the downgrading of the practitioners of medicine in the social hierarchy, by the law-givers, and also to interpolations of theological/philosophical discussions in the *Caraka-samhitā*, which are inconsistent with its rational spirit.

The conflict between science and the counter-ideology was also present in

Europe in the middle ages. There was the inquisition, the burning at the stake of nonconformist thinkers like Giordano Bruno, the trial and humiliation of Galileo, and the indexing of books. How does one compare the conflict between science and the counter-ideology in India with that in Europe? Why was it that the counter-ideology could prevail in India, whereas Europe could get over its influence?

Is the argument of Chattopadhyaya regarding suppression of materialism in relation to Indian medicine overdrawn? It would be useful to have a fresh look at it. Medicine did thrive in India, did make advances[27] and even spread outside India down to the eighth century. Indian doctors were imported by the Abbasid Khalifas to supervise their own hospitals in the 8th century AD and Sanskrit medical texts were translated into Arabic[28].

A Spanish Muslim scholar, quoted by Chattopadhyaya himself, says in 1080 AD that Indian medicine was the best in the world.

Further, when Francis Buchanan carried out a survey of the state of medicine and surgery in 1807, under orders from the Court of Directors of the East India Company, he observed in relation to Bihar: '*Medicine is taught by several of the Pandits, some of whom also, though they are grammarians, practice the art . . . Besides the Professors of Medicine, about 700 families of Brahmans, almost all of Sakadwip, practice that art and are the only Hindu physicians who possess anything like science . . .*'[29] (emphasis added).

Conflict between science and religion

It is not unusual to find it claimed that unlike in Europe, there was no conflict between science and religion in India[30, 31], because there was no hostility to science or to new ideas in philosophy, on the part of religion; that even atheists were not persecuted in India. This was true in early times, when there was full intellectual freedom, a beautiful example of openness being the *Nāsadāya sūkta* in the *Rigveda*. There were as many as sixty-three systems of philosophy in the days of the Buddha, and very few of them cared to rely on Vedic authority for their hypotheses or their conclusions. Within the fold of the orthodoxy itself there were the *Sānkhya*, *Mimāmsa* and *Nyāya* systems, which had hardly any appropriate place for the divine Creator in them. Buddhism and Jainism were not summarily dismissed as atheistic. But certain works came to be canonized some time around 600 BC, and an almost equal reverence came to be paid to the *Smrtis* and *Puranas* in the course of time. Theories began to be accepted or rejected according as they were in conformity with or opposed to statements of the sacred books[32]. Already by the time of Sankara, the pressure was quite strong. Intellectual giants like Sankara and Rāmānuja had to spend a disproportionate amount of time and energy to prove that their systems of philosophy were in conformity with and the natural outcomes of the *Upanisadic* hypotheses. Independent development of philosophies was affected. Instead, one only had commentaries on the revealed literature[32].

The separation of philosophy from theology, and of science from philosophy, also did not take place. This separation would seem to have been an important step in European development. It enabled science to go ahead when theology and metaphysics stagnated.

Under such circumstances, there was not much scope left for research and originality in those matters where opinions were expressed in the sacred books, which now included *Smrtis* and *Purānas*. A glaring example is that of the mythological stories about *Rāhu* and *Ketu* temporarily overpowering the moon or the sun and thereby causing eclipses, which got canonized by inclusion in the *Purānas*. Aryabhata, (*ca.* 500 AD), Varahamihira and Brahmagupta (*ca.* 600 AD) knew the scientific explanation of eclipses but were powerless to explode the myths. In fact, Brahmagupta went to the extent of condemning Aryabhata, Varahamihira, Srisena and Visnucandra for expounding the unorthodox but scientific theory of eclipses. Alberuni has suggested that Brahmagupta made the compromise because a calamitous fate like that of Socrates had befallen him[33].

Was this also the reason why Aryabhata's suggestion that the earth revolves around its axis was not followed up by later astronomers? Or, for the fact, that the idea of impetus could not develop into inertia, force, etc., – perhaps the most important developments in Europe that launched the scientific revolution?

Incidentally, it will be useful to have more light thrown on the process of canonization of scriptures, including the *Smrtis* and the *Purānas*, which made it impossible for even outstanding intellectuals like Brahmagupta to counter stories like those of *Rāhu*, *Ketu* and the eclipses, and also to document other instances of this kind. Is the view that there was no conflict between science and religion/spirituality a modern view, or does it go back to ancient times?

Swami Ranganathananda[34] quotes from Sankarāchārya's commentary on the *Brhadāranyaka Upanisad*: 'Nor are the *Srtis* (the *Vedas*) supposed to have authority in matters which are contradicted by other means of knowledge, for instance, if they said, 'fire is cold and wets things'. Was such thinking widely accepted by scholars or people at large? During which epochs?

Possibly connected with this concept of knowledge as captured and frozen in the scriptures is the picture of a traditional '*Brahmin*' scholar today as a repository of knowledge in the books – rather than a free-wheeling creative intellectual, applying his knowledge to the exploration of new phenomena or providing improved explanations of the phenomena already known. When did the tradition of questioning as in the *Upanisads*, *Nyāya-Vaisesika*, and other early texts lose its vigour and a static view of the learning of the learned emerge?

The belief that all knowledge is in the Vedas

Vivekananda refers to the belief prevalent in India that all that is called knowledge is in the *Vedas*[35].

It is not uncommon to find it claimed even today that the *Vedas* are the storehouse of all knowledge – *spiritual and secular*. For example, Jagadguru Swami Shri Bharati Krishna Maharaj has said:

'The very word *Veda* has this derivational meaning, i.e. the fountainhead and illimitable storehouse of all knowledge. This derivation, in effect, means, connotes and *implies that the Vedas should contain within themselves all the knowledge needed by mankind relating not only to the so-called*

spiritual (or other-worldly) matters but also to those usually described as purely secular, temporal or worldly, and also to the means required by humanity as such for the achievement of all-round, complete and perfect success in all conceivable directions, and that there can be no adjectival or restrictive epithet calculated (or tending) to limit that knowledge down in any sphere, any direction or any respect whatsoever.

'In other words, it connotes and implies that our ancient Indian Vedic lore should be all-round, complete and perfect and be able to throw the fullest necessary light on all matters which any inspiring seeker after knowledge can possibly seek to be enlightened on.'[36] (emphasis added)

One should note that this statement of the Jagadguru is quoted by a former President of the Association of Indian Universities in his Foreword to the book *Science and the Vedas*.

It will be useful to know how old this kind of belief is in our cultural/philosophical history. When and how did it get ascendancy? Such a belief, if widely held, could lead to stagnation or even decline in science and could, therefore, be an important factor inhibiting the growth of modern science in India. It could also have inhibited the healthy development of philosophy.

Tendency to accommodate conflicting opposites

Indian tendency to accommodate conflicting opposites has been mentioned as one of the causes. The faith of some of our modern scientists in miracle-making godmen is cited as a present-day example.

Alberuni, the scholar–traveller who visited India around 1000 AD was very critical of the Hindus on this score: '. . . you mostly find that even the so-called scientific theorems of the Hindus are in a state of utter confusion, devoid of any logical order, and in the last instance always mixed up with silly notions of the crowd, e.g. immense numbers, enormous spaces of time, and all kinds of religious dogmas, which the vulgar belief does not admit of being called into question. Therefore, it is a prevailing practice among the Hindus *jurare in verba magistri*; and I can only compare their mathematical and astronomical literature, as far as I know it, to a mixture of pearl shells and sour dates, or of pearls and dung, or of costly crystals and common pebbles. Both kinds of things are equal in their eyes, since they cannot raise themselves to the methods of a strictly scientific deduction . . . '[37]

These are harsh words, possibly coloured to some extent by the arrogance of a person accompanying conquering Mahmud, but they should make us look introspectively at the cultural and philosophical milieu of that period.

Varāhamihira wrote in *Brahatsamhita* that the gems could have originated either from the bones of gods and demons, or on account of natural metamorphosis of rocks. The possibilities were given equal emphasis[38]. Biswas[38] has given more revolting examples from *Rasa–Ratna–Samuccaya*, a 13th century text.

Kosambi[39] has observed: 'The "logic" advanced by the *Brahmins* took good care to avoid all reality. The end result is seen in the philosophy of the great Sankara (800 AD), who threw out the proposition that "a thing is either A or non-A", and viewed the universe as divided into metaphysical categories upon several planes. The highest plane was, of course, of speculation about and unity with eternal principles. Material reality did not exist. The philosopher was thus excused if he joined the common herd on the plane of ritual observation . . . '

An oft-quoted example of this tendency as far as science is concerned is that Brahmagupta wrote the popular myth about the cause of eclipses (*Rāhu*, *Ketu* and all that) along with his calculation in terms of the correct astronomical explanation.

It is interesting to see what Saha[40] has said in this regard: '. . . in India, the votaries of the astronomical science apparently found it difficult to go against the scriptures. The Greek savants, with characteristic boldness and freedom of thought, had pooh-poohed the Homeric and Hesiodic cosmogony, they made light of the Olympic Gods. But the Indian savant never showed equal boldness. On the other hand, being accustomed to easy-going pantheistic tendencies, he was always ready to compromise and be accommodating in his views . . . never courageous enough to propose a bolder theory of the universe in opposition to current ideas. Aryabhata of Kusumpur (Patliputra) (born in 476 AD) thought that the earth, by turning round its axis, caused day and night, but his successors did not seem to be impressed with his ideas. It is, however, clear that they were keenly alive to the stupidity of public beliefs or utter inadequacy of the *śāstric* lore. "It is said", says Bhāskaracharya, "that the earth rests on the hood of a huge snake, which when it feels uneasy gives a shake and causes earthquakes. Well, the snake must have something to rest upon, let it be the tortoise, but this also must have something else to rest upon. So we have to assume the hypothesis of an endless string of the support and supporters. A better explanation would be that the earth has no support, i.e. is suspended in space". But the next sentence, he is careful to say, "but since these things are mentioned in the *Sastras*, there must be some truth in them"'.

Lest we are too harsh on Brahmagupta or Bhaskaracharya, we may note what Gould[41] has to say about Newton and Whiston, his successor as Lucasian Professor at Cambridge.

'The Newton of our Pantheon is a sanitized and modern version of the man himself.' He greatly admired the book *New Theory* by William Whiston (1696), to such an extent that Whiston ended up as Newton's handpicked successor as Lucasian Professor of Mathematics at Cambridge. The full title of Whiston's treatise gives an idea of its contents. Whiston called his treatise *A New Theory of the Earth from its Origin to the Consummation of all Things, Wherein the Creation of the World in Six Days, the Universal Deluge, and the General Conflagration, as laid down in the Holy Scriptures, are shown to be perfectly agreeable to Reason and Philosophy.*

In trying to establish a consistency between the two great sources of truth as defined by his countrymen – the infallibility of the Scriptures, and the mathematical beauty of the cosmos, so recently recreated by Newton – Whiston implicated comets as prime movers.

'Whiston has descended through history', remarks Gould[41], 'as the worst example of religious superstition viewed as an impediment to science'. '. . . We dismiss Whiston because he violated ideals of science as we now define the term. But, in Whiston's time, science did not exist as a separate domain of inquiry; the word itself had not been coined. No matter how we may view such an enterprise today, Whiston's mixture of natural events and scriptural traditions defined a primary domain at

the forefront of scholarship in his time . . . He did not view his effort as a work of science, but as a treatise in an important contemporary tradition for using all domains of knowledge – revelations of the Scriptures, history of ancient chronicles, and knowledge of nature's laws – to reconstruct the story of human life on our planet. The *New Theory* contains – and by Whiston's explicit design – far more material on theological principles and biblical exegesis than on anything that would now pass muster as science.'

The interesting thing in the context of our discussion is the fact that Whiston's work was highly thought of by no less a person than Newton.

It is also well known that at the same time that Newton was working on the *Principia*, he was also spending a substantial fraction of his time on investigations in alchemy and theology. On the other hand, he did not publish anything on either of these subjects.

Complacency developed during the period when the Indians had attained a high level of civilization

Another line of reasoning puts the blame on the complacency that the Indians had developed over a period of time: they began to think that they had attained the highest knowledge that humans were capable of and they progressively lost the questioning attitude and openness to knowledge from other sources (*ā no bhadrāh kratavo yantu visvatah*). In this context one often quotes Alberuni, who says the following about the attitude of Hindus of his time:

'They are haughty, foolish, vain, stolid and self-conceited. According to their belief, there is no other country on earth but theirs, no other race of man but theirs, and no created beings besides them have any knowledge or science whatsoever. Their haughtiness is such that, if you tell them of any science or scholar in Khurasan or Persis, they will think you to be both an ignoramus and a liar. *If they travelled and mixed with other nations, they would soon change their mind, for their ancestors were not as narrow-minded as the present generation is.*'[42] (emphasis added).

Gargi Samhita[43] says: 'The yavanas (i.e. the Greeks) are indeed barbarians, but astronomy originated in them and for this they must be venerated as Gods.'

Aryabhata and Varāhamihira also did not have the attitude described by Alberuni, though one probably sees the beginning of such an arrogant and exclusive attitude in *Varahamihir* when he says: 'The Greeks, *though impure*, must be honoured, since they were trained in sciences, and therein excelled others. *What then are we to say of a Brahman if he combines with his purity the height of science*?'[44] (emphasis added).

When did the vain, self-conceited attitude start dominating? How did it continue in spite of the humiliating defeats that the Hindu society had to accept from successive invaders?

The complacency or intellectual lethargy that had developed continued even after the Moguls came to power. Jawaharlal Nehru refers to the curious fact that Akbar did not take any interest in building sea power, though Vasco de Gama had reached Calicut, via the Cape, in 1498, and established Portuguese sea power in the Indian Ocean. What is more, this lack of interest continued even after the Portuguese started extracting tolls from traders and from pilgrims going to Mecca. He or the 'jewels' in his court also took no interest in printing[45], though the Jesuits presented him with a printed Bible. The Mogul army and the armies of other states in India depended on foreign experts for their artilleries. Why did Akbar or anyone else not send his own men abroad for training, or interest himself in the improvement of the artillery by encouraging research work? Why did they not take interest in developing Indian clocks, though imported clocks were very popular with the Mogul nobility?

Over the centuries, during the Gupta period and later, India imported horses from Arabia, Iran and Bactria. For some strange reason, India never took to the breeding of horses and training the veterinaries. Mentioning this, Thapar[46] remarks: 'The fact that India never bred sufficient horses of quality, the best blood having always to be imported, was to have disastrous consequences on the cavalry arm of Indian armies, eventually making the cavalry ineffective, especially in comparison with Central Asian hordes.'

Rigidity of the caste system

A factor that needs some discussion is the growing rigidity of the caste system in India, which led to the separation of the head from the hand[47]: Theoretical and philosophical studies (including the speculative aspects of science) were the prerogative of the *Brahmins*, whereas the practical arts and crafts were relegated to the other castes. These lower castes, moreover, did not have the benefit of formal education. This prevented the coordination between science and crafts, and theory and experimentation/experience. Also, the accumulated theoretical knowledge was accessible to *Brahmins* in Sanskrit while the common technicians, who knew only one of the Prakrits, had no access to it. It would be useful to compare and contrast with a similar problem in medieval Europe, where the accumulated theoretical knowledge was the prerogative of the Latin-knowing clergy and aristocracy and not accessible to common artisans. Also, when did the caste system become rigid from the point of view of separation of the head from hand? Altekar observes that after 800 AD, primary education became restricted to the higher strata of society; the vast majority of workmen and artisans became illiterate by this time and knew only the traditional processes of their profession[48].

The ban on foreign travel

Foreign travel seems to have got banned around the 8th century[49]. This is usually attributed to the increasing rigidity of the caste system, and the ideas of purity and impurity, especially in relation to mixing with foreigners and observing food taboos. Vivekananda in one discussion attributes the ban to a desire to prevent the Hindus from mixing with the surrounding Buddhistic nations. He says: 'The one great cause of the downfall and degeneration of India was the building of a wall of custom – whose foundation was the hatred of others – around the nation, and the real aim of which in ancient times was to prevent the Hindus from coming into contact with the surrounding Buddhistic nations.'[51]

Whatever be the cause, banning of foreign travel could not but have had a very adverse impact on the exposure to fresh ideas and new techniques, so essential for growth of science. Sharma[52] has discussed the possibility that Raja Jai Singh may have received wrong advice from his foreign consultants, the Jesuits, with regard to the state-of-the-

art in astronomy in Europe. Though their delegation visited Europe on his behalf, the Jesuits, because of their own theological beliefs and prejudices, may not have presented the Copernican revolution to Jai Singh in its true perspective. The European books which Jai Singh got translated into Sanskrit included the already obsolete *Almagest* of Ptolemy and not *The Revolution of the Heavenly Spheres* of Copernicus or *The Harmonies of the World* of Kepler or *Dialogues Concerning Two New Sciences* of Galileo or *The Principia* of Newton, all of which had been in existence for quite some time. Jai Singh also does not seem to have benefited from the developments in telescopes taking place in Europe, in planning his observatories. 'Unfortunately', says Sharma, 'The Raja seemed to have had no alternative but to send a team of foreigners on his fact-finding mission in Europe as no high-born Hindu of his days would dare '*cross the ocean*' and thereby risk 'losing his caste'[52]. This unfortunate predicament of the Raja need not diminish our admiration for his vision and his contributions to astronomy.

A handicap of the oral tradition?

The Indian tradition is said to have been largely an oral tradition, from the earliest times when the *Vedas* were not written but passed on by recitation from the *Guru* to the disciple, with explanations in the course of the transfer. Later when writing was developed (*ca.* 400 BC), the style was usually aphoristic, so that the intermediary of a *Guru* was needed for an understanding of the meaning of the esoteric texts. While commentaries helped to some extent, they often did not throw light on the rationale of the problem tackled or provide an insight into the mental working of the scientist, whose results were given cryptically in a *Sūtra*. Added to this was the secretive aspect of the tradition, which again, meant that if the symbolic language with secret meaning, or recipé, or special formula, was not transferred by a *Guru* to his (favourite) *Śisya*, that piece of knowledge disappeared[53].

Effect of invasions?

Successive invasions meant not only devastating destabilization of the functioning of the society, and loss of prosperity, it also meant loss of centres of learning, including their libraries. Alberuni[42] describes the effect of Sultan Mahmud's invasions: '. . . Mahmud utterly ruined the prosperity of the country, and performed there wonderful exploits, by which the Hindus became like atoms of dust, scattered in all directions, and like a tale of old in the mouth of people . . . This is the reason why Hindu sciences have retired far away from those parts of the country conquered by us, and have fled to places which our hand cannot yet reach . . . '.

India had at one time flourishing universities at Taksaśila, Vikramaśila, Valabhi, Nālandā, etc. Within a couple of hundred years of the above account by Alberuni, the 700–800 years old University of Nālandā, to which scholars came from various parts of the country and from abroad, was destroyed by Bakhtiyar Khilji. Bakhtiyar Khilji also destroyed the famous University of Vikramasila. These were the last of the great Indian universities to disappear. No great universities came up in the centuries that followed, either under the Afghans and the Moguls in the North, or under the Vijayanagar and other kingdoms of the South.

One may say that one had the *Gurukul* system. But such a *gharānā* system had its weak as well as strong points. It could not be a substitute for universities.

It would not be an accident that the bright period for Indian science and philosophy was also the period in which some of these universities flourished and attracted scholars from far-off places. One notes in this context the role played by the universities of Cordova, Cracow, Tübingen, Paris, Oxford, Bologna and Padova in the revival of learning (science in particular) in Europe. Some of the great architects of the scientific revolution in Europe had been at Padova at important periods in their career – Copernicus, Galileo, Harvey and Vesalius, to mention a few.

Even the great Akbar does not seem to have applied his mind to the setting up of universities. Why?

Just around the time when India's great universities were disappearing, new universities were coming up in Europe[54]. This was followed by the setting up of Academies/Royal Societies, which also were to play an important role in the intellectual life of Europe.

The effect of centuries of unsettled conditions that prevailed before a multi-religious composite culture could emerge needs to be analysed dispassionately without allowing either Hindu or Muslim chauvinism to distort the view of the happenings.

The British had stepped in before India could find its feet again.

Influence of Buddhism?

Going back to an earlier period, what were the positive and negative impacts of Buddhism on the development of science in India? The Buddha himself stressed the importance of the spirit of enquiry:

'Believe nothing
Merely because you have been told it,
Or because it is traditional,
Or because you yourself have imagined it.
Do not believe what your teacher tells you,
Merely out of respect for the teacher.
But whatever, after due examination and analysis,
You find to be conducive to the good,
The benefit,
The welfare of all beings,
That doctrine believe and cling to,
And take it as your guide.'

It is interesting that he said this around the time that some of the Hindu scriptures were getting canonized. Can one trace the healthy impact of this on science?

Later, things changed. As Chattopadhyaya points out in the course of his discussion of the vicissitudes of *Ayurveda*, while the concept of causality is absolutely crucial for *Ayurveda*, Nāgārjuna, the founder/exponent of the *Mādhyamaka* school of philosophy, popularly known as *Śūnyavada*, denied causality and condemned experience and reason as sources of right knowledge. What impact did this have on science?

It is curious that the decline of Indian science seems to have started around the time or in the centuries following the victory of *Ādi Śankara* over the existing Buddhistic doctrines and other materialistic and nonmaterialistic schools of Indian philosophy in the 8th century. Did this victory imply the reestablishment of spirit over matter, lead-

ing to a down-grading of creative science[55]?

Did the *Prāmānyavāda*, which followed in the succeeding centuries, have an adverse effect?

One hopes that an objective discussion of the positive and negative impacts will be forthcoming[56].

Too big a load from the past?

Was the continuity in Indian civilization a handicap? Europe lost its Graeco-Roman knowledge and tradition in the dark ages, and then recovered these a thousand years later through the Arabs, along with the knowledge which the Arabs brought from India. When this happened, the Europeans had a chance to have a fresh critical look at the old heritage. Such a critical look started already in the 13th–14th centuries with the scholastics, and was an important factor leading to Renaissance. Was the absence of discontinuity a disadvantage for the Chinese and the Indians?

The load from the past has been summarized by Kosambi: 'The absence of logic, contempt for mundane reality, the inability to work at manual and menial tasks, emphasis upon learning basic formulas by rote, with the secret meaning to be expounded by a high *Guru*, and respect for tradition (no matter how silly) backed by fictitious ancient authority had a devastating effect upon Indian science.'[57]

The various corrosive factors seem to have been operating from the last centuries of the first millennium, though their effect was seen much later. As Nehru[58] has remarked, during Akbar's time, 'India to all outward seeming presented the spectacle of not only a brilliant court, but of thriving trade, commerce, manufacture and crafts. In many respects, the countries of Europe would have seemed backward and rather crude to an Indian visitor then. And yet the dynamic quality which was becoming evident in Europe was almost wholly absent in India . . . A civilization decays much more from inner failure than from external attack. It may fail because in a sense it has worked itself out and has nothing more to offer to a changing world, or because the people who represent it deteriorate in quality and can no longer support the burden worthily . . . '.

Some comments on the western traditions supposedly responsible for the rise of modern science in Europe

Finally, it may be appropriate to comment very briefly on some of the contributory factors that are often mentioned in the context of the rise of modern science in Europe.

Rise of modern science was a unique happening, that took place in Europe. There has been a lot of post-facto rationalization of this development[59]. The search has been for developments which took place uniquely in Europe, so as to make a case that modern science could arise only as a concomitant of such developments and, therefore, in Europe and nowhere else. Developments like Reformation and Renaissance are often mentioned as contributory or triggering factors. It is good to remember, however, that Copernicus was not only a catholic but a canon in the Roman Catholic Church and that high dignitaries of the Church were associated with the publication of his book. Galileo was also a Roman Catholic and ventured into the realm of theology by attempting to show that the Copernican system was consistent with a fair interpretation of the words of the Scriptures. And Martin Luther mounted a vicious attack on Copernicus. The Catholic Church, it is well known, put the books of Copernicus and Galileo on the *Index*.

Anthony[60], a biographer of Newton, quotes Butterfield to make the point that the ancient Greek thought, through Arab translations, Renaissance, Reformation, etc., did not change the mode of thinking of people as scientific revolution itself did.

Butterfield has, in fact, observed, 'since that (i.e. scientific) revolution overturned the authority in science not only of the middle ages but also of the ancient world – since it ended not only in the eclipse of scholastic philosophy but also in the destruction of Aristotlean physics – it outshines everything since the rise of Christianity and reduces the Renaissance and Reformation to the rank of mere episodes, mere internal displacements within the system of medieval Christiandom . . . '[61].

The question then arises if one has been overemphasizing factors like Renaissance, Reformation and changes in the mode of thinking of people as the prime contributors to the rise of modern science in Europe.

And Bertrand Russell has observed:

'Although science was what interested (Francis) Bacon, and although his general outlook was scientific, he missed most of what was being done in science in his day. He rejected Copernican theory', even after the work of Kepler. 'Bacon appears not to have known of the work of Vesalius, the pioneer of modern anatomy, or of Gilbert, whose work on magnetism brilliantly illustrated the inductive method. Still more surprising, he seemed unconscious of the work of Harvey, although Harvey was his medical attendant!'[62].

It is also interesting to note the comments of Thomas Huxley on the influence of Baconian philosophy:

' . . . I do protest that, of the vast number of cants in this world, there are none, to my mind, so contemptible as the pseudo-scientific cant which is talked about the Baconian philosophy. To hear people talk about the great chancellor – and a very great man he certainly was – you would think that it was he who invented science, and that there was no such thing as sound reasoning before the time of queen Elizabeth . . . The method of scientific enquiry (did not originate with Bacon or his age) originated with the first man whoever he was, . . . '[63].

Butterfield[64] has commented that one of the aspects of history that is not sufficiently emphasized is that '*from the fourth to the twentieth century, one of the most remarkable aspects of the story – the most impressive conflict that spans 1500 years – is the conflict between Europe and Asia, a conflict in which down to the time of Newton's Principia, it was the Asiatics who were on the aggressive*'.

It is clear that some of the contributory factors held responsible for the rise of modern science in the West and for the East lagging behind need a fresh look.

Mukhopadhyaya[65] has written about two metamorphoses of the mind in the recent centuries, one among the Europeans, who appropriated science as a purely western product, as a most typical creation of thought of the Western civilization, and the other suffered by the colonial people, Indians in particular, which persuaded them that science was not a part of their civilization.

Mukhopadhyaya's thesis calls for a closer look.

Concluding remarks

It should be clear that the Project on the History of Indian Science, Philosophy and Culture should constantly keep in view the changing sociocultural–philosophical milieu in which science could have flourished in India during certain periods, and in which Indian technology could have been at the forefront internationally in certain domains down to the eighteenth century, and in which both became stagnant or declined during other periods. What were the external factors and what were the internal failures *which must not be allowed to repeat*? Were the internal failures on a broad front (science, philosophy, literature, art, architecture, . . .)? To the extent that the Project will succeed in unravelling these complex intertwining factors, it will provide a landmark in the history of the development of Indian science, philosophy and culture.

1. Chattopadhyaya, D. P., On the nature of the interconnection between science, technology, philosophy and culture, Occasional Paper 1 of the PHISPC; also the Project Document.
2. Seal, B. N., *Positive Sciences of the Ancient Hindus*, Motilal Banarasidas, Delhi, 1958.
3. Chattopadhyaya, Debi Prasad, *The History of Science and Technology in India: The Beginnings*, Firma KLM, Calcutta, 1986.
4. Sen, S. N., History of science in relation to philosophy and culture in Indian civilization, Occasional Paper 6 of the PHISPC.
5. Bose, D. M., Sen, S. N. and Subbarayappa, B. V. (eds.), *A Concise History of Science in India*, Indian National Science Academy, New Delhi, 1971.
6. Dharampal, *Indian Science and Technology in the Eighteenth Century: Some Contemporary European Accounts*, Impex India, New Delhi, 1971.
7. Alvarez, Claude, *Homo Faber*, Allied Publishers, Bombay, 1979.
8. Udgaonkar, B. M., Scientific culture and ideological influences on history of science in India, Occasional Paper 16 of the PHISPC.
9. *Mundaka Upanisad*, I.1.1–2.
10. Joshi, Kireet, Significance of Indian yoga (an overview), Occasional Paper 10 of the PHISPC. (It would be useful for Kireet Joshi to discuss the impact of such a thinking on the development of science.)
11. Related to this is the fact that in India, where mythology continues to play an important role in people's daily life, stories of mythological (and today's) personalities acquiring supernatural knowledge, skills, powers by *sadhana* continue to be believed by many people. We even have the phenomenon of a Prime Minister addressing an academic gathering and making a reference to our ancients having visited *Candraloka*.
12. Nehru, Jawaharlal, *The Discovery of India*, John Day and Co., New York, 1946, p. 180.
13. Thapar, Romila, *A History of India*, Penguin Books, Harmondsworth, England, 1966 (Reprinted 1981), p. 115.
14. Bernier, François, quoted by B. V. Subbarayappa in ref. 5, p. 485.
15. Swami Vivekananda, *Jnana-Yoga*, 15th edn, Advaita Ashrama, Calcutta, 1980, p. 47.
16. Swami Vivekananda, *Jnana-Yoga* 15th edn, Advaita Ashrama, Calcutta, 1980, pp. 91–92.
17. Swami Vivekananda, *Jnana-Yoga* 15th edn, Advaita Ashrama, Calcutta, 1980, p. 93.
18. Swami Vivekananda, *Jnana-Yoga* 15th edn, Advaita Ashrama, Calcutta, 1980, pp. 98–99.
19. Swami Vivekananda, *Jnana-Yoga* 15th edn, Advaita Ashrama, Calcutta, 1980, p. 49.
20. Thapar, Romila, *The Past and Prejudice*, National Book Trust, 1975 (Reprint 1990), p. 7.
21. Nehru, Jawaharlal, *The Discovery of India*, John Day and Co., New York, 1946, pp. 72–73.
22. Thapar, Romila, *The Past and Prejudice*, National Book Trust, 1975 (Reprint 1990), pp. 18–19.
23. Alvarez, Claude, *Homo Faber*, Allied Publishers, Bombay, 1979, p. 48.
24. Chattopadhyaya, Debi Prasad, *Science and Society in Ancient India*, Research India Publication, Calcutta, 1977.
25. Chattopadhyaya, Debi Prasad, Science, philosophy and society in ancient India, Paper presented at the XVIth International Congress of the History of Science, Bucharest, August 26–September 3, 1981.
26. Chattopadhyaya, Debi Prasad, *Science and Society in Ancient India*, Research India Publication, Calcutta, 1977.
27. Unfortunately, soon after Susruta's time, dissection of human bodies went out of vogue, to the great detriment of the progress of medical science (ref. 28, p. 178).
28. Altekar, A. S., *Education in Ancient India*, The Indian Bookshop, Banaras, 1934, pp. 180–181.
29. Sen, S. N., *Scientific and Technical Education in India 1781–1900*, Indian National Science Academy, New Delhi, 1992, pp. 26–27.
30. Swami Ranganathananda, *Science and Spirituality*, Ramakrishna Math, Hyderabad, 1982; also *Science and Religion*, Advaita Ashram, Calcutta, 1982.
31. It is not our intention to belittle the valuable efforts being made by Swami Ranganathananda to build bridges between religion and spirituality on the one hand, and science on the other, in the context of modern needs. The question, however, is what happened in history.
32. Altekar, A. S., *Education in Ancient India*, The Indian Bookshop, Banaras, 1934, pp.342–346. Altekar also discusses the narrowness that developed in the educational system over the centuries.
33. Alberuni, in *Alberuni's India*, (ed. Sachau, Edward), Popular Edition, 1914, vol. II, p. 112.
34. Swami Ranganathananda, *Science and Spirituality*, Ramakrishna Math, Hyderabad, 1982.
35. Swami Vivekananda, *Jnana-Yoga*, 15th edn, Advaita Ashrama, Calcutta, 1980, p. 166.
36. Jagadguru Swami Bharati Krishna Maharaj, quoted in the Editor's Foreword to *Sciences and the Vedas* (ed. Patel, Ishwarbhai, Former President, Association of Indian Universities), Somaiya Publications, 1984.
37. Alberuni, in *Alberuni's India* (ed. Edward Sachan), Popular Edition, 1914, vol. I, p. 25.
38. Biswas, A. K., *Indian J. Hist. Sci.*, 1987, **22**(1), 29.
39. Kosambi, D. D., *The Culture and Civilization of Ancient India in Historical Outline*, Vikas, New Delhi, 1977, p. 163.
40. Saha, M. N., *Collected Works of Meghnad Saha* (ed. Chatterjee, Santimoy), Orient Longman, Bombay, 1982, vol. 1, pp. 83–84.
41. Gould, Stephen Jay, in *Bully for Brontosaurus*, Penguin, 1992, pp. 367–381.
42. Swami Ranganathananda, *Science and Spirituality*, Ramakrishna Math, Hyderabad, 1982; also *Science and Religion*, Advaita Ashram, Calcutta, 1982, vol. 1, p. 22. Note, however, that Alberuni has elsewhere praised the Hindus and their contributions to the science. He even learnt Sanskrit and translated some of the works.
43. Quoted by Majumdar, R. C., in *History and Culture of Indian People*, Bharatiya Vidya Bhavan, New Delhi, vol. II, Chapter XXIV.
44. Varahamihir, quoted by Alberuni in Ref. 33, pp. 22–23.
45. Daniel, J. Boorstein has drawn attention to the fact that 'despite an uncanny reverence for the written word, the Arabic world resisted the opportunities of the printing press', and remarks that 'fear of blasphemy and heterodoxy kept the printing press out of the Muslim world for centuries'. See Boorstein Daniel, J., *The Discoverers*, Vintage Books, 1985, pp. 539–547.

46. Thapar, Romila, *A History of India, loc. cit*, p. 149.
47. Bose, D. M., Sen, S. N. and Subbarayappa, B. V. (eds.), *A Concise History of Science in India*, Indian National Science Academy, New Delhi, 1971, pp. 484–485.
48. Altekar, A. S., *Education in Ancient India*, The Indian Bookshop, Banaras, 1934, pp. 198.
49. How effective was the ban until the 13th century? Until then, Indians had close interaction with West/Central Asia and SE Asia[50]. And it could not have affected Muslim scholars and traders. The Gujaratis were, in fact, the dominant ship-owning and trading community at Malacca at the end of the fifteenth century (see, for example, *The Suma Oriental of Tomé Pires*, Hakluyt Society, London, 1944, 2 volumes, translated and edited by Armando Cortesão, quoted in *South-East Asia: Its Historical Development*, by John F. Cady, Tata McGraw Hill, New Delhi, 1976).
50. Habib, Irfan, in *Science and Technology in Indian Culture – A Historical Perspective*, NISTADS, New Delhi, 1984.
51. Swami *Vivekenanda*, *Teachings of Swami Vivekananda*, Advaita Ashrama, Calcutta, 1971, 5th edn, p. 154.
52. Sharma, Virendra N, *Indian J. Hist. Sci.*, 1982, **18**, 333, 345.
53. It is a pleasure to thank B. V. Subbarayappa for a discussion on the contents of this paragraph.
54. It is interesting to read the announcement of the founding of a new University at Toulouse in 1229, which made a special mention that 'those who wish to scrutinize the bosom of nature to the inmost can hear here the books of Aristotle, which were forbidden at Paris'. Ref. 'New University-1229' in *A Random Walk in Science*, Institute of Physics, London, 1973, p. 15; the item being condensed from Thorndike, Lynn, *University Records of the Middle Ages*, Columbia University Press, 1944.
55. Ray, P. C. said in this context: 'The *Vedanta* Philosophy, as modified and expounded by Sankara, which teaches the unreality of the material world, is to a large extent responsible for bringing the study of physical science into disrepute. Sankara is unsparing in his stricture on *Kanada* and his system ... Among the people riden by caste and injunctions of the *Vedas*, *Puranas* and *Smrtis* and the like, cramped and paralysed, no Boyle could arise.' (quoted by Debi Prasad Chattopadhyaya in Ref. 3)
56. See Sen, S. N., History of science in relation to philosophy and culture in Indian civilization, Occasional Paper 6 of the PHISPC, for a brief discussion.
57. Kosambi, D. D., *The Culture and Civilization of Ancient India in Historical Outline*, Vikas, New Delhi, 1977, pp. 174–175.
58. Nehru, Jawaharlal, *The Discovery of India*, John Day and Co., New York, 1946, p. 261
59. Singh, Virendra, *Asia-Pacific Phys. News*, 1987, 2, 20–24, for a recent review and several references.
60. Anthony, H. D., *Sir Isaac Newton*, Abelard Schuman, London, 1960.
61. Butterfield, H., *The Origins of Modern Science*, Collier Books, New York, 1962, pp. 7–8.
62. Russell, B., *A History of Western Philosophy*, Simon and Schuster, p. 544.
63. Huxley, Thomas Henry, Six Lectures to Working Men, 'On our knowledge of the causes of phenomena of organic nature', 1863, reprinted in *Darwinia*.
64. Butterfield, H., *The Origins of Modern Science*, *loc. cit.*, p. 189.
65. Mukhopadhyaya, P. K., History of science and two metamorphoses of mind, Occasional Paper 9 of the PHISPC.

ACKNOWLEDGEMENT. I thank Dr B. V. Subbarayappa for a critical reading of the manuscript.

B. M. Udgaonkar is in Homi Bhabha Centre for Science Education, Tata Institute of Fundamental Research, Mankhurd, Bombay 400 088, India.

Edited and published by Prof. S. Ramaseshan and Prof. P. Balaram, Current Science Association, Bangalore 560 080.
Typeset by Creative Literati Pvt. Ltd., Bangalore (Tel. cum Fax: 5462698) Printed at Printek, Bangalore (3357763)

Council of Scientific & Industrial Research (CSIR), New Delhi, India

Advt. No. 6/95

Position: Director, National Aerospace Laboratories (NAL), Bangalore

CSIR, the premier agency established by the Government of India to undertake scientific and industrial research in the country, is looking for a suitable scientist/technologist for the post of Director for its National Aerospace Laboratories (NAL), Bangalore. NAL is devoted to carrying out application-oriented R&D work in aerospace engineering and related areas with particular reference to design, construction and operation of aircraft, maintenance and operation of important testing and analytical facilities and undertakes contract R&D work in its areas of activity.

The post carries the pay scale of Rs 5900–200–7300 plus allowances as admissible to Central Government officials. Consultancy remuneration. The appointment to the post will be on contract for a period of six years or superannuation, whichever is earlier.

For further details, please contact the Director-General, CSIR, Rafi Marg, New Delhi 100 001 (India), (Fax: 3710618, Telex: 031–65202 CSIR IN, 031–66147 CSIR IN, Gram: CONSEARCH, NEW DELHI).

Council of Scientific & Industrial Research (CSIR), New Delhi, India

Advt. No. 5/95

Position: Director, National Chemical Laboratory (NCL), Pune

CSIR, the premier agency established by the Government of India to undertake scientific and industrial research in the country, is looking for a suitable scientist/technologist for the post of Director for its National Chemical Laboratory (NCL), Pune. The Laboratory is devoted to carrying out research in different areas of chemistry, developing new technologies for gainful utilization of natural resources and assisting the chemical industry through consultancy, contract R&D and training.

The post carries the pay scale of Rs 5900–200–7300 plus allowances as admissible to Central Government officials. Consultancy is permitted without any ceiling on the annual consultancy remuneration. The appointment to the post will be on contract for a period of six years or superannuation, whichever is earlier.

For further details, please contact the Director–General, CSIR, Rafi Marg, New Delhi – 110 001 (INDIA), (Fax: 3710618, Telex: 031–65202 CSIR IN, 031–66147 CSIR IN, Gram: CONSEARCH, NEW DELHI).

Institute for Research in Reproduction (ICMR)

The Institute for Research in Reproduction is happy to announce a short-term training course on Techniques in Immunology, Cell Biology and Molecular Biology from 2 November 1995. The training course will cover the latest techniques in these areas.

Applications are invited from mid-career scientists/technologists (preferably below 40 years) working on regular positions in universities, national laboratories, research institutions and in-house R&D centres.

The candidates should apply giving details in the following format:

(i) Educational qualification, research experience, age
(ii) Present place and nature of work
(iii) Facilities for R&D available in candidate's laboratory
(iv) Present and future research interest of the candidate's laboratory
(v) How the candidate proposes to make use of the training in his/her research programmes.

About 12 candidates will be selected. The participants will not be charged registration fees. However, TA/DA will have to be met by the sponsoring institutions. Accommodation for outstation candidates will be made available at subsidized rates.

Applications duly sponsored and through proper channel should reach the Director/Convenor, Dr (Mrs) U. C. Hegde, Institute for Research in Reproduction, Jehangir Merwanji Street, Parel, Bombay 400 012 (Fax: 91–022–4139412) on or before 31 August 1995. Selected candidates will be informed by 15 September 1995.

Nuclear Science Centre

Inter-University Research Facility
(An Autonomous UGC Institution)

Advt. 3/95

Applications are invited for the following:

SCIENTIST TRAINEE (3 positions: one general, others one each reserved for SC & ST): M Sc in Physics, persons with M Sc before 1993 need not apply. Fixed stipend of Rs 2500 during the training period with free hostel-type accommodation for one year. If successful after training, may be placed in detector/cryogenics laboratory (grade Rs 2200-4000).

SCIENTIST (Atomic Physics), Grade Rs 3000–4500: Ph D in Experimental Atomic Physics, minimum 2 years of experience in accelerator-based research. Meritorious candidates with relevant experience may be considered for higher grade.

ENGINEER (Cryogenics), Grade Rs 3000–4500: B E (Mech.) or M Sc (Physics), minimum 4 years of experience in cryogenics. Hands-on experience on He refrigeration system is required. Familiarity with cryostats/gas handling system and computer-aided designing will be considered as added qualification.

ENGINEER (Experimental Facility Development), Scale Rs 3000–4500: B Tech (Instrumentation/ Electronics) after B Sc (Physics) or M Sc (Tech/Applied Physics) or M Sc (Physics with Electronics), minimum 4 years of experience in development of instrumentation and computer interfacing. Experience in both hardware and software is preferred.

Send curriculum vitae with complete details with photograph and attested marksheets from High School onwards and names of three referees. Information should reach **A.O. (Admn.), Nuclear Science Centre, Aruna Asaf Ali Marg, Post Box 10502, New Delhi 110 067**, within 30 days from the day of this advertisement.

CURRENT SCIENCE

A fortnightly journal of research

Editors: S. Ramaseshan and P. Balaram

Editorial Board

S. Arunachalam, Central Electrochemical Research Institute, Karaikudi 623 006
Asis Datta, School of Life Sciences, Jawaharlal Nehru University, New Delhi 110 067
D. Balasubramanian, Centre for Cellular and Molecular Biology, Hyderabad 500 007
J. Chandrasekhar, Department of Organic Chemistry, Indian Institute of Science, Bangalore 560 012
R. Gadagkar, Centre for Ecological Sciences, Indian Institute of Science, Bangalore 560 012
K. N. Ganeshaiah, Department of Plant Genetics and Breeding, University of Agricultural Sciences, Bangalore 560 065
R. V. Hosur, Tata Institute of Fundamental Research, Bombay 400 005
N. V. Joshi, Centre for Ecological Sciences, Indian Institute of Science, Bangalore 560 012
Jyotirmoy Das, Indian Institute of Chemical Biology, Calcutta 700 032
C. C. Kartha, Division of Cellular and Molecular Cardiology, Sree Chitra Tirunal Institute for Medical Sciences and Technology, Thiruvananthapuram 695 011
S. S. Krishnamurthy, Inorganic and Physical Chemistry Department, Indian Institute of Science, Bangalore 560 012
S. C. Lakhotia, Department of Zoology, Banaras Hindu University, Varanasi 221 005
K. Muralidhar, Department of Zoology, Delhi University, Delhi 110 007
R. Narasimhan, CMC Limited, Bangalore 560 001
K. Neelakantan, ANURAG (Advanced Numerical Research and Analysis Group), Hyderabad 500 258
R. Nityananda, Raman Research Institute, Bangalore 560 080
T. Padmanabhan, Inter-University Centre for Astronomy and Astrophysics, Pune 411 007
B. R. Pai, Propulsion Division, National Aerospace Laboratories, Bangalore 560 017
T. J. Pandian, School of Biological Sciences, Madurai Kamaraj University, Madurai 625 021
Y. S. Rajan, Department of Science and Technology, New Delhi 110 016
M. Ramakrishnan, Geological Survey of India, Calcutta 700 016
A. V. Ramani, T. T. K. Pharma Ltd., Bangalore 560 025
S. Ramaswamy, Department of Physics, Indian Institute of Science, Bangalore 560 012
A. K. Raychaudhuri, Department of Physics, Indian Institute of Science, Bangalore 560 012
S. R. Shetye, Physical Oceanography Division, National Institute of Oceanography, Dona Paula, Goa 403 004
V. Siddhartha, 51 Bharati Nagar, New Delhi 110 003
V. Srinivas, School of Mathematics, Tata Institute of Fundamental Research, Bombay 400 005
G. Srinivasan, Raman Research Institute, Bangalore 560 080
R. Uma Shaanker, Department of Crop Physiology, University of Agricultural Sciences, Bangalore 560 065
K. S. Valdiya, Department of Geology, Kumaun University, Nainital 263 002
G. Venkataraman, Department of Physics, Sri Sathya Sai Institute of Higher Learning, Prasanthi Nilayam 515 134
M. Vidyasagar, Centre for Artificial Intelligence and Robotics, Bangalore 560 001
K. VijayRaghavan, TIFR Centre, Indian Institute of Science Campus, Bangalore 560 012

Current Science, founded in 1932, is published by the Current Science Association in collaboration with the Indian Academy of Sciences. The journal is also intended as a medium for communication and discussion of important issues that concern science and scientific activity. All articles published in *Current Science*, especially editorials, opinions and commentaries, letters and book reviews, are deemed to reflect the individual views of the authors and not the official points of view, either of the Current Science Association or of the Indian Academy of Sciences.

Working Committee, Current Science Association

P. Balaram, Indian Institute of Science, Bangalore
R. Chidambaram, Department of Atomic Energy, Bombay
V. K. Gaur, CSIR Centre for Mathematical Modelling and Computer Simulation, Bangalore
K. Kasturirangan, ISRO Satellite Centre, Bangalore
P. K. Kaw, Institute for Plasma Research, Gandhinagar
N. Kumar, Raman Research Institute, Bangalore (*Vice-President*)
N. V. Madhusudana, Raman Research Institute, Bangalore (*Secretary*)
A. Mani, Field Research Unit, Indian Institute of Tropical Meteorology, Bangalore (*Past-President*)
R. A. Mashelkar, National Chemical Laboratory, Pune
R. Narasimha, Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore
G. Padmanaban, Indian Institute of Science, Bangalore (*Vice-President*)
R. Ramachandran, The Institute of Mathematical Sciences, Madras
P. Rama Rao, Department of Science and Technology, New Delhi
S. Ramaseshan, Raman Research Institute, Bangalore (*Treasurer*)
M. S. Valiathan, Manipal Academy of Higher Education, Manipal (*President*)

Editorial Office

Editorial staff

Chandrika Ramesh, G. Madhavan, C. S. Ravi Kumar, M. S. Venugopal

Circulation and Accounts

Peter Jayaraj, Ranjini Mohan, B. Sethumani, Shanthi Bhaskar, B. K. Shivaramiah

Annual subscription 1995

	Institutional	Personal
India	Rs 350	Rs 100
All other countries (AIR MAIL)	US $175	US $50

Single copies other than special issues: Rs 35/US $15

Editorial Office: CURRENT SCIENCE
C. V. Raman Avenue
P.B. 8001
Bangalore 560 080, India
Telephone: 91-80-3342310
Telex 845-2178 ACAD IN
Fax: 91-80-3346094

Copyright © 1995, Indian Academy of Sciences

GOVERNMENT OF INDIA
MINISTRY OF SCIENCE & TECHNOLOGY
DEPARTMENT OF SCIENCE & TECHNOLOGY

PROPOSALS INVITED FROM INDUSTRIES FOR PARTIAL FINANCIAL SUPPORT FOR DEVELOPMENT OF INSTRUMENTS

The Department of Science & Technology (DST) under its scheme on 'Instrument Development Programme' (IDP) has been promoting and supporting the programmes on indigenous development of sophisticated instruments. The Department has been providing financial support to R&D organizations for design and development of identified instruments in the areas of Analytical, Medical, Geo-scientific, Agri-electronic and Test and measuring instruments. Industries (private or public sectors) are associated with the projects from the very beginning to ensure commercial exploitation of the technology.

The Department now extends the programme inviting suitable project proposals from industrial firms having well-established in-house R&D units recognized by DSIR for partial financial support for development of identified instruments in the following areas:

1. Analytical instruments
2. Medical instruments
3. Geo-scientific instruments
4. Agri-electronic instruments
5. Test and measuring instruments
6. Laser-based instruments
7. Food processing instruments
8. Environment monitoring and pollution control instruments; and
9. Sensors

The individual firms should have good track record of R&D achievements for seeking financial support. Preference will be given to joint project proposals from industries and R&D organizations. However, the Department reserves the right to reject any proposal without assigning any reason thereof. For more information/guidelines about submission of project proposals, the interested parties may write to:

Joint Adviser (IDP)
Instrument Development Division
Department of Science & Technology
Ministry of Science & Technology
Technology Bhawan, New Mehrauli Road
New Delhi 110 016
Tel: 011-662463, 667373 (Extn. 267)
Fax: 011-661682, 6862418, 6863847

Founded: 1932

ISSN: 0011-3891

CURRENT SCIENCE

Volume 69 Number 3

10 August 1995

211 In this issue

CORRESPONDENCE

213 **Revitalization of Indian science**
S. C. Tiwari

NEWS

215 **1994 Kalinga Prize ■ New appointments**

RESEARCH NEWS

216 **Oxide electronics**
C. N. R. Rao

217 **Noncentrosymmetric monolayers from centrosymmetric molecules – A fresh approach to materials for frequency doubling**
T. P. Radhakrishnan

219 **Targeting RNA in the cell**
S. C. Lakhotia

OPINION

221 **Declining phase of quality research in our universities**
Samir Bhattacharya

COMMENTARY

223 **Implications of recent levelling observations for Tehri and other high dams in the Himalaya**
R. Chander and V. K. Gahalaut

SCIENTIFIC CORRESPONDENCE

227 **Storing river runoff in the Gulf of Khambhat: A critique of the 'Kalpasar' concept**
S. K. Gupta and P. Sharma

228 ***In situ* preserved *Vertebraria* axes in Ib River Coalfield, Orissa**
Kamal Jeet Singh and Shaila Chandra

229 **Role of phospholipase C in the cryptobiotic cysts of the fairy shrimp *Streptocephalus dichotomus***
G. K. Shanmugasundaram and N. Munuswamy

GENERAL ARTICLES

231 **Reflecting diffusions**
S. Ramasubramanian

237 **Structural, clinical and transmission aspects of dioxins – Potential environmental pollutants**
C. Ganesh Kumar and K. Jayaraj Rao

240 **The birth of the semiconductor superlattice**
L. Esaki

RESEARCH ACCOUNT

243 **NMR studies on the Myb-DNA system**
P. K. Radha, Anup Madan, L. C. Padhy and R. V. Hosur

RESEARCH ARTICLE

252 **Lithium and tungsten mineralization in Sewariya pluton, south Delhi fold belt, Rajasthan: Evidences for preferential host rock affinity**
Subrata Banerji and M. K. Pandit

RESEARCH COMMUNICATIONS

257 **Chemotherapy reduces human anti-filarial IgE response**
M. K. Beuria and Manoj K. Das

258 **Biochemical basis for the differentiation of the two nonpoisonous snakes *Eryx conicus* Schneider and *Eryx johnii* Russell (family: Boidae)**
P. Senthil Kumar, A. P. Kamalakara Rao and T. M. Vatsala

263 **Effect of addition of SO_2 on solid-state fermentation of apple pomace**
V. K. Joshi, D. K. Sandhu and S. Jaiswal

265 **Analysis and significance of female reciprocal call in frogs**
Debjani Roy, Bijoylakshmi Borah and Amarendra Sarma

270 **Yttrium-, niobium- and zirconium-rich rhyolite dykes of Dhorio Nes area, district Jamnagar, Gujarat, India**
P. B. Maithani, Rahul Banerjee, Navin Goyal, S. Ramachandran and Rajendra Singh

272 **Evaluation of efficacy of bricks as geochemical monitor of atmospheric heavy-metal pollution**
Rohit Shrivastav, Sanjay Kumar Mathur, Shobhit Shrivastav, M. M. Shrivastav, Sahab Das and Satya Prakash

BOOK REVIEWS

276 **Mahatma Gandhi: An Apostle of Applied Human Ecology,** *reviewed by* M. S. Swaminathan

277 **Recent Trends in Aerobiology, Allergy and Immunology,** *reviewed by* S. V. Gangal

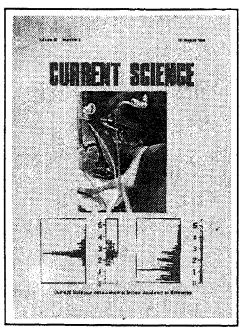
277 ***In vitro* Culture and its Applications in Horticulture,** *reviewed by* R. S. Nadgauda

HISTORICAL COMMENTARY AND NOTES

279 **Two undocumented historic earthquakes and additional information on the major earthquake of 1 September 1803 from north India**
R. N. Athavale

ERRATUM

278 **Testosterone metabolism by rhesus monkey spermatozoa: Effects of antifertility agents**
M. Rajalakshmi, R. S. Sharma, S. Mokkapati and P. C. Pal (*Curr. Sci.*, 1995, **68**, 1121)



COVER. *Top*: Male and female *Rana erythraea*. Male on the left and female on the right. *Bottom*: *left*, Mean spectrum and sonagram of the male advertisement call of *Rana erythraea*; *right*, Mean spectrum and sonagram of the female reciprocal call of *Rana erythraea*. See page 265.

Indexed in CURRENT CONTENTS/GEOBASE/CHEMICAL ABSTRACTS

In this issue

What can we learn from the croak of a frog?

For the common man, frogs bring on the image of slimy, creepy, ugly things that croak without a trace of decency. For the Indian undergraduate student of zoology, frogs must appear as having been created by God for the sole purpose of being used for teaching and learning how to dissect animals and study their anatomy. But like the princess in the fairy tale who kissed a frog and turned him into a prince, Debjani Roy and her students might well bring to the frog a large measure of respect and admiration.

On **page 265** of this issue, Debjani Roy, Bijoylakshmi Borah and Amarendra Sarma describe experiments that permit a deeper understanding of the role of frog calls in their courtship behaviour. Because of the feebleness of their calls and perhaps at least in part because of the gender of most previous researchers on frogs, the calls of the female frogs have largely been neglected. Most previous work has concentrated on male calls and has created the view that males advertise themselves so that females have no difficulty in finding them. The present study has made a special attempt to go beyond the din created by the male frogs and attempt to listen to what the females have to say. The recordings and analysis of the female calls suggest that females are calling in response to the males and in the context of courtship. Typical of males in general, the male frogs call as loud as possible in an attempt to get the attention of the females. On the other hand, the females' feeble call is probably testing the sensitivity of the males!

In any case this study suggests a much more active role for the female frog in matters of mate choice and courtship than has been hitherto suspected. This is reminiscent of the more general shift we are now witnessing, from a view of animal courtship based largely on male–male competition for access to the passive females towards a view where female choice has a significant role in deciding who mates with whom. That somebody has thought it worthwhile to leave the laboratory bench and the dissection tray and go to the inhospitable marshes at night, and that too in North East India, to learn from living frogs, augurs well for Indian zoology. There is so much to learn from the croak of a frog.

R.G.

NMR: Fifty years on

Almost fifty years have passed since Felix Bloch and Edward Purcell independently reported the observation of nuclear magnetic resonance (NMR) in bulk matter. Conceived as an esoteric technique to measure nuclear magnetic moments accurately, hardly anyone in the late 1940s could have foreseen the impact that NMR was to make in the fields of chemistry, biology and medicine. The first sign that NMR was too useful to be left entirely to physicists came early on; the proton resonance frequencies in the samples used by Bloch (tap water) and Purcell (mineral oil) were different. Clearly, chemistry was at work, modifying Larmor precession frequencies. A more dramatic manifestation of the role of chemical environment in influencing nuclear resonance frequencies was provided in 1951 with the famous demonstration of three chemically shifted groups of proton resonances in ethyl alcohol by Arnold, Dharmatti and Packard. The last named was to go on to developing commercial NMR instrumentation, while Dharmatti was to introduce the NMR technique at the then fledgling Tata Institute of Fundamental Research (TIFR) in Bombay. Interestingly, the first magnetic resonance experiments in India were conceived and executed by G. Suryan in the Physics Department of the Indian Institute of Science, Bangalore, in 1950–51. By the early 1960s NMR was a critically important tool for structure determination in organic chemistry, with the hydrogen nucleus as the sole reported case of chemical environment. No chemistry department in the West would consider itself complete if it did not possess an NMR spectrometer (a condition that has now been reached in India also). Two problems, however, limited the range of NMR applications; sensitivity of detection of resonances and inadequate chemical shift dispersion in complex molecules containing large numbers of atoms. The introduction of Fourier Transform Spectroscopy in the late 1960s and the advent of superconducting magnets capable of providing high magnetic fields of requisite homogeneity allowed NMR to breach the barriers between chemistry and biology. Enhanced sensitivity now permitted detection of resonances from less abundant nuclei (^{13}C, ^{15}N), enhancing the information content of NMR studies. But more was to come. In the late 1970s following a suggestion by Jean Jeneer, the first two-dimensional NMR experiments were performed by Richard Ernst. Overlapping frequencies could now be resolved by spreading the information onto a frequency plane; more importantly correlations between coupled nuclei permitted identification of groups of spins characteristic of specific chemical groups. By the early 1980s there was a veritable flood of multiple pulse experiments; spectroscopists could now literally eavesdrop on intimate conversations between atomic nuclei in molecules. The convenient detection of nuclear Overhäuser effects (NOEs) permitted identification of spatially proximate (< 4 Å) hydrogen atoms, providing finally a firm handle for the determination of three-dimensional structures of molecules in solution. NMR was poised by the mid 1980s to enter an arena which had till then been exclusively dominated by X-ray crystallography: the field of biological structure determination.

The first pioneering study of bovine pancreatic trypsin inhibitor, a small 58-residue polypeptide by Kurt Wüthrich and his collaborators at the ETH in Zurich, was quickly followed by a deluge. The size range of proteins that can be studied by NMR methods has steadily increased with the incorporation of stable isotopes (^{13}C, ^{15}N), permitting multidimensional spectroscopy. Advances in NMR have paralleled the revolution in molecular biology, which permits ready production of recombinant proteins; the fusion of the two areas now promises to provide remarkably detailed information on biological structures in solution.

In India too, the sophistication of the available NMR methodology has slowly increased with the TIFR now housing spectrometers that operate at 500 and 600 MHz for proton resonances. In this issue (**page 243**) P. K. Radha *et al.* describe a structural study of the DNA binding domain of a proto-oncogene product, c-Myb, from *Drosophila*. The analysis of the NMR spectrum of the 160 residue polypeptide and the determination of the structure of a dodecameric nucleotide containing the segment recognized by the c-Myb protein using multidimensional NMR methods are described.

P.B.

INDIAN INSTITUTE OF SCIENCE
BANGALORE 560 012

Applications are invited for a FACULTY POSITION in the Department of Physics at the level of **ASSISTANT PROFESSOR** in the scale of Rs 3700–125–4950–150–5700.

Qualifications and Experience

Should have a Ph D in Physics with first class in the preceding degree; a good record of publications and 3 years of postdoctoral teaching/research experience (relaxable in exceptional cases) in the field of optical physics. The candidate should have specialized in the broad area of experimental optical physics and nonlinear optics.

He/she is expected to make optimal use of the major existing facilities in a well-equipped laser laboratory of the Department.

Details and application form can be obtained by sending a written request accompanied by an IPO for Rs 10/- drawn in favour of the Registrar (free of cost to SC/ST candidates on production of caste certificates) along with a self-addressed stamped (Rs 2/-) envelope of size 28 × 18 cm on or before 30 August 1995.

Completed applications along with copies of certificates, publications and a research plan for the next few years should reach the Registrar on or before 15 September 1995. Referees may be requested to send their considered reports to the Registrar.

R(IA)308-14/95
Dated: 26 July 1995

REGISTRAR

CURRENT SCIENCE

Volume 69 Number 3 10 August 1995

CORRESPONDENCE

Revitalization of Indian science

A draft paper[1] prepared by the Indian Academy of Sciences on science education, National Science University (NSU) proposal of Mahajan[2], and the dedication[3] to the nation of the Jawaharlal Nehru Centre for Advanced Scientific Research (JNCASR) on 4 March 1995 have not only commonality in the declared objectives of achieving excellence in science, they also display identical mind-sets of the persons involved. My purpose is two-fold: to delineate some of the vexed questions related with the Indian science by analysing these examples[1-3], and to offer suggestions for small corrective measures to check the degeneration effectively.

The debate on the NSU proposal has been mainly centred on three issues discussed below:

1. *The Role of the NSU.* The basic premise is that a talented few should be nurtured by providing them with the support and facilities comparable to the best in the world. This is not new, and as noted by many readers the so-called elite institutions have been established based on this premise without having made much impact on the state of science done in our country. As for catching up and competing with the frontline research areas, the path-breaking or sufficiently influential contributions from India have been almost nil. If we examine the names of the scientists comprising the top body in any heavily funded enterprise ostensibly aimed at excellence, then we find almost the same set of persons with minor changes here and there. Obviously, the policy makers believe that there are only a handful of such high-calibre competent scientists. As a corollary, if the NSU comes into being, the honorary faculty of the JNCASR or such other 'eminent' scientists will be appointed in that also. Therefore, assuming that ideally whatever is proposed in Mahajan's scheme gets implemented, there will be marginal impact on Indian science, there will be no depletion of faculty from the universities, and for ordinary scientists it will be irrelevant. Why should we be alarmed? The JNCASR was established without any open discussion. Why? Is the NSU proposal being criticized because the group of powerful scientists perceive this as a threat to their monopoly? One talks of NRI arrogance, but have the Institute scientists ever realized their own arrogant behaviour towards the University and college teachers? The Academy paper[1] approves the recommendations made by the Planning Commission Working Group for a three-tier scheme. Tier I is aimed at 0.5% of the estimated total UG science students in a year, with the allocation of funds of the order of 22% of the total grant. And in tier III 83% of the students will also get almost the same amount, i.e. 22%. I cannot imagine if any idea can be worse than this, and fully endorse Sriram's rejection[4] of this 'few highly talented' syndrome.

2. *Criticism of science establishment.* Many readers have felt offended by sweeping condemnation of feudal science establishment (or coterie) by Mahajan. While it is true that the language used by Mahajan is unpleasant (or pedestrian according to Murthy[5]), the criticisms are valid. The Academy paper[1] makes a general statement without any study: 'In most universities and affiliated colleges, one finds low educational standards and a poor academic environment.' According to Sriram[4]: 'In fact, scientists in our country miss no opportunity to declare that our universities have low standards and poor academic environment.' Not only commoners do not need resources, they can be denied basic dignity! From my experience I have found that rapid decline in teaching has occurred primarily due to misplaced unimaginative policies and schemes formulated by eminent scientists, i.e. centres of excellence, frontline research, and promotions linked with Ph D. Ironically, such policies have been lauded in ref. 1. In the centres of excellence, costly equipment are inaccessible to the students and teachers, the personnel appointed are mediocre and pliable, the so-called frontline research activities pursued are done in isolation within the faculty, and control of funds in few hands leads to exclusivity and negative impact on the general teaching environment.

Returning to the existence of coterie, I shall give a recent example of powerful coterie in action. On 16 February 1995, during a Discussion Meeting of the JNCASR at Banaras Hindu University, C. N. R. Rao delivered an evening talk on Indian science, in which he was harshly critical of CSIR laboratories, their Directors and Bhatnagar Prize winners. Although several of them were present, none expressed even mild resentment. Interaction with students and teachers of the eminent scientists was cosmetic. The whole exercise seemed to be a demonstration of the powerful coterie.

In my opinion the crucial factor responsible for the creation of this feudalism in science is the exalted position granted to the science administrators close to the political power, and lack of transparency in the policy matters. For example, Arunachalam[6] has made very pertinent comments on the election of the Secretary, Department of Science and Technology (DST), as the new President of the Academy.

3. *Patriotism.* Swadesh Mahajan is a nonresident scientist in the US, so why not attack his patriotism? Many comments have appeared[5,7] implying that the concern shown for India by NRIs is of

doubtful nature; their hidden aim is to promote their self-interest. I think NRIs are not necessarily opportunists, and those working in India are not necessarily inspired by nationalism – we must recognize this. It is better to remember that for C. V. Raman, J. C. Bose and M. N. Saha patriotism was an article of faith which came to them naturally. Bhabha essentially introduced the concept of what I call contractual nationalism in science. Most of the scientists who have returned to India have done so after prolonged bargains. And the patriotism of quite a few top scientists is so shaky that slight dilution in their status or facilities and funding makes them repent their stay in our country. I refer to an interview[8] and quote C. N. R. Rao's reply: 'Probably I should have left India long ago'.

Science establishment

Science education cannot be discussed in isolation; support and funding for science and technology cannot be divorced from the socioeconomic conditions of the people on the one hand and obligations of the scientists towards the society on the other. For a serious discussion, obsession with excellence, and undue importance given to science administrators and managers must go. Excellence in the midst of common people emerges naturally, and it can be easily identified if the intentions are honest. Raman demonstrated this in his life, and fortunately his views on 'centre of excellence' have also been recorded by Ramaseshan[9]. I quote a few sentences: '... I strongly believe that fundamental science cannot be driven by instructional, industrial, governmental or military pressures .... Among the many qualities called for in a person who assumes this responsibility are scientific integrity, vision, receptiveness to new ideas and the enlightened outlook to let younger people grow unhindered to their full height...'

The scientists in top positions have shown ideological elasticity to suit political leadership of the day. To take a very important example, at the Prime Minister's initiative in 1976, Indian Science Congress focused attention on rural development. Since then started the convention of focal themes. The idea as such is quite good. But what has happened to this? The task forces were established, and recommendations for follow up actions were made[10]. I list some of them: '... Social responsibilities of scientists; Assessment and improvement of biofertilizers utilizing genetic engineering techniques wherever possible and the possibility of their utilization on a large scale in agriculture; Protection of tribals and weaker groups from exploitation of their natural resources; Universities and institutions should be adequately funded in order to (a) revise the standard of scientific education, (b) provide long- and short-term training for young scientists to deal with socioeconomic problems, and (c) operate extension wings to enable laboratory findings to reach the rural areas; Engineering expertise should be pooled and tied up with prime plant manufacturers for standardization of plants and equipment, which may reduce lead time and optimize investment.' These are very significant and relevant suggestions for the country, but the science establishment has not shown any enthusiasm and commitment to implement them. Today economic liberalization has led top scientists to find virtues in privatization of higher education, university–industry interaction, marketism of R&D, and shift from rural people to limiting 'tertiary education' for selected few. The conclusion is that not only have they (i.e. top scientists) failed, they cannot be trusted for an independent and fair stand.

It has been advised that the science administrators who have failed to deliver the goods should retire[11], and the demand for a white paper has been made by C. N. R. Rao[12]. Both the suggestions deserve to be implemented.

Corrective measures

Attention is drawn to the suggestions made earlier[13] with some clarifications here. The reason for the abolition of centres of excellence in universities has already been elaborated in the present article. However, in case of doubts their performance and the psychological impact of these centres on the rest of the faculty may be evaluated in open and transparent manner. The newly set up National Accreditation and Assessment Council of the UGC may look into this matter. I shall briefly state few more corrective measures. No person after retirement should be appointed on any administrative, Government's policy-making body or in the selection committees. The system of students' union and teachers' association in the universities and colleges should be abolished; instead, a scheme for redressal and a vigilance council may be formulated. The hierarchical system of teaching positions should be discontinued. To check inefficiency and corruption in the administration, a computerized system and reduction of present ministerial staff to less than 25% should be considered.

Role of academies

In a recent editorial candid observations about the Academy have been made by Balaram[14], and first report on higher education by the Indian Academy of Sciences has been published[1]. It is interesting to note that in UK a group called National Academies Policy Advisory Group was established in 1992, and the first report has come out on Intellectual Property and the Academic Community[15]. Whether our Academy has been inspired by that is not important, the step taken is most welcome. S. W. Hawking was elected FRS in 1974 'while holding the lowly position of research assistant at Cambridge'[16]. Perhaps the Indian Academies need to be made broad-based, and to learn from such examples for recognizing excellence. Arunachalam's comment[6] merits serious consideration. I think as a policy the office bearers of the Academies should not be chosen from amongst those holding administrative posts in the government.

In my opinion, the role of Academies in revitalizing Indian science and education can be very profound. Awareness, action and accountability could be the focus of attention.

Awareness. Information dissemination on policy matters, and forum for discussion and debate on the issues. *Current Science* has done praiseworthy beginning in this direction. The editorial[17] has made explicit this policy; however, I may remark that even now only less than 10% of the space is devoted to this aspect (e.g. less than 90 pages out of 976 pages in volume 66 (1994)), which includes all categories of articles published on the policy matters. Further the reach of the journal is limited; therefore, a cheap volume of reprint collection of the articles published during the last 3–4 years on this topic may be published by the Academy for students and teachers in

colleges. A scheme whereby the subscribers of *Current Science* may be motivated to gift one year subscription to someone in colleges may be initiated.

Action. The Academies can do certain things without depending on the government policies. For example, making available small grants for writing books to college teachers, and setting up of a national teachers scheme. In the latter case, the fellows selected as national teachers shall be required to spend a minimum of one month in a college, and three months in an academic year at different colleges. During their stay, formal lectures and informal interaction with the students and teachers will be their main duty.

Accountability. Corruption in various forms has become rampant in research and teaching. The Academies can play a very effective role in dealing with this problem. Scientists and Academics found guilty in plagiarism or other frauds[18] can be blacklisted: their work will not be published, their fellowships or titles awarded by the academies shall be withdrawn, and their names along with their malpractices shall be made public.

Conclusion

We have enough strength and capability, which need to be properly utilized; there are many creative minds, they only need to be identified. The sense of belonging and involvement in the system by all ordinary scientists and Academics should be cultivated. Frustration, helplessness and servile mentality must go; obviously, hard and unpleasant (for few) actions may be necessary for that to happen.

1. University Education in Science, Academy report, *Curr. Sci.*, 1995, **68**, 255–267.
2. Mahajan, S. M., *Curr. Sci.*, 1994, **67**, 503–508.
3. Mukunda, N., *Curr. Sci.*, 1995, **68**, 570–572.
4. Sriram, M. S., *Curr. Sci.*, 1995, **68**, 235–236.
5. Murthy, T. K. S., *Curr. Sci.*, 1995, **68**, 7–9.
6. Arunachalam, S., *Curr. Sci.*, 1995, **68**, 23–24.
7. Nityananda, R., *Curr. Sci.*, 1994, **67**, 767–768; Vidyasagar, M., *Curr. Sci.*, 1994, **67**, 769–770; Jayakrishnan, A., *Curr. Sci.*, 1995, **68**, 9–10.
8. Rao, C. N. R., *Frontline*, 12 March 1993.
9. Ramaseshan, S., *Prajna*, 1987–1988, 123–124.
10. Consolidated Report on the Recommendations of the Indian Science Congress Session 1976–1980, Department of Science & Technology, New Delhi.
11. Narasimhan, N. S., *Curr. Sci.*, 1994, **67**, 772–773.
12. Rao, C. N. R., *The Times of India*, 5 August 1993.
13. Tiwari, S. C., *Curr. Sci.*, 1994, **66**, 10–14.
14. Balaram, P., *Curr. Sci.*, 1995, **68**, 233.
15. *Nature*, 1995, **374**, 398.
16. *Current Contents*, 1995, **35**, 3–7.
17. Ramaseshan, S. and Balaram, P., *Curr. Sci.*, 1995, **68**, 5.
18. Gupta, Y. P., *The Hindustan Times*, 21 February 1995; Paintal, A. S., *Curr. Sci.*, 1994, **66**, 905–907.

S. C. TIWARI

Department of Physics,
Banaras Hindu University,
Varanasi 221 005, India.

NEWS

1994 Kalinga Prize

The 1994 Kalinga Prize for the popularization of science has been awarded to Nikolai Nikolaevich Drozdov, Associate Professor of Biogeography at the Moscow State University.

Established by UNESCO in 1951, it is an annual award of £1000 based on a grant to UNESCO from B. Patnaik of the state of Orissa, India, the founder and President of Kalinga Foundation Trust. With the 1994 awardee, the tally now stands at 49 personalities from 19 countries (Jain, N. C., *Curr. Sci.*, 1994, **66**, 618–620).

A graduate of Moscow University in 1963, Dr Drozdov has done his doctorate in 1968 on the ornithofauna of the deserts of the USSR.

While working at the Australian National University (Canberra), he published a book called *The Flight of the Boomerang* about his journey through Australia, about animals, nature and nature conservation. It instantly became a best-seller and over 100,000 copies of its two editions published in 1980 and 1988 were sold out within a couple of weeks of its release. For the last twenty-three years Dr Drozdov has been presenting a bimonthly television programme called *In the World of Animals.* This highly popular one-hour programme on animal life in the vast CIS territory is watched by approximately 200 million viewers. In 1992 he won the **Golden Panda** for outstanding achievement at the Wildscreen Festival in Bristol. In 1989 he was included in Global 500, a list drawn up by the United Nations Environment Programme of the world's leading ecologists and specialists in environmental protection.

New appointments

Dr V. S. Ramamurthy, 53, has taken over as Secretary of the Department of Science and Technology in place of Dr P. Rama Rao on 1 July. Ramamurthy, a nuclear physicist, was earlier the Director of the Institute of Physics, Bhubaneswar.

Dr R. A. Mashelkar, 52, has taken over as Director-General of the Council of Scientific and Industrial Research in place of Dr S. K. Joshi on 1 July. Mashelkar, a chemical engineer, was the Director of the National Chemical Laboratory at Pune.

Oxide electronics

C. N. R. Rao

Metal oxides have emerged as promising electronic materials of the 21st century.

Silicon occupies a preeminent position in electronics today and this unique status of Si will probably remain for many years to come. Compound semiconductors such as GaAs were considered to be good candidates to replace Si because of their versatility, but this is far from happening. Yet, one is tempted to ask the question 'Will there be a material or a class of materials as good as or superior to silicon?'. Considering the wide range of properties exhibited by them, metal oxides would appear to be the likely contenders. The nature of bonding in oxides can vary anywhere between ionic and metallic or covalent. Oxides also exhibit a phenomenal range of electronic and magnetic properties[1,2]. Thus, we have oxides with metallic properties (e.g. RuO_2, ReO_3, $LaNiO_3$) at one end of the range and oxides with highly insulating behaviour (e.g. $BaTiO_3$) at the other. There are also oxides that traverse both these regimes with change in temperature, pressure or composition (e.g. V_2O_3, $La_{1-x}Sr_xCoO_3$). Interesting electronic properties also arise from charge density waves (e.g. $K_{0.3}MoO_3$), charge ordering (e.g. Fe_3O_4) and defect ordering (e.g. $Ca_2Mn_2O_5$, $Ca_2Fe_2O_5$). Oxides with diverse magnetic properties anywhere from ferromagnetism (e.g. CrO_2, $SrRuO_3$, $La_{0.5}Sr_{0.5}MnO_3$) to antiferromagnetism (e.g. NiO, $LaCrO_3$) are known. Many oxides possess switchable orientation states as in ferroelectric (e.g. $BaTiO_3$, $KNbO_3$) and ferroelastic (e.g. $Gd_2(MoO_4)_3$) materials. No discovery in solid-state science has created as much sensation, however, as that of high-temperature superconductivity in cuprates[3]. Although superconductivity in transition metal oxides has been known for some time, the highest T_c reached was around 13 K; we now have oxides with T_c's in the region of 160 K, which is not far from the lowest temperature recorded on earth (183 K). The discovery of high-T_c cuprates has focused worldwide attention on the physics, chemistry and materials applications of metal oxides.

Many applications of oxide materials in electronics are already known. They range from the use of oxides as insulators, capacitors, piezo-electric and pyroelectric materials, sensor materials, varistors, solid ionics, transparent conductors, and in magnetic and optical devices. In such applications, oxides are used in the form of bulk solids or thin films, generally in polycrystalline form. Fine ceramic technology is based on the control of grain size on the micrometer scale. Use of nanoparticles of oxidic materials is also gaining importance. With the progress in microelectronics, there has been much interest in the fabrication of thin films, especially in the epitaxial (high-quality single crystal) phase. A current concern in the development of large integrated silicon electronics is focused on high dielectric constant oxide films for capacitors and nonvolatile memories. In this context, heteroepitaxial growth of perovskite-type oxide films on silicon substrates is of great interest. Extensive investigations are being conducted for this purpose by employing various methods such as vacuum evaporation, sputtering, chemical vapour deposition (CVD), sol-gel process, pulsed laser deposition, and molecular beam epitaxy (MBE). Films fabricated hitherto are by no means exotic and possess crystal structures of known perovskites, possibly with minor modifications.

The phenomenon of giant magnetoresistance (GMR), that was known earlier in metallic bilayers and magnetic semiconductors, has recently been found to occur in oxides as well. In the last year, there has been intense research on perovskite oxides of the type $La_{1-x}A_xMnO_3$ (A = alkaline earth) exhibiting GMR[4]. GMR is found in bulk oxides as well as in films and this property will clearly be exploited for several applications.

One of the highlights of advanced electronics research today relates to artificially designed structures to test device concepts based on quantum effects (quantum dots, wires and wells). For this purpose, investigations on atomic control and characterization of heterostructures are underway in many laboratories. These structures are all based on Si and related materials. Until recently, such atomically engineered oxides were not explored and were considered impossible. The situation has, however, changed since the discovery of high-T_c superconductors. The potential of oxides as advanced electronic materials has enhanced enormously because of the

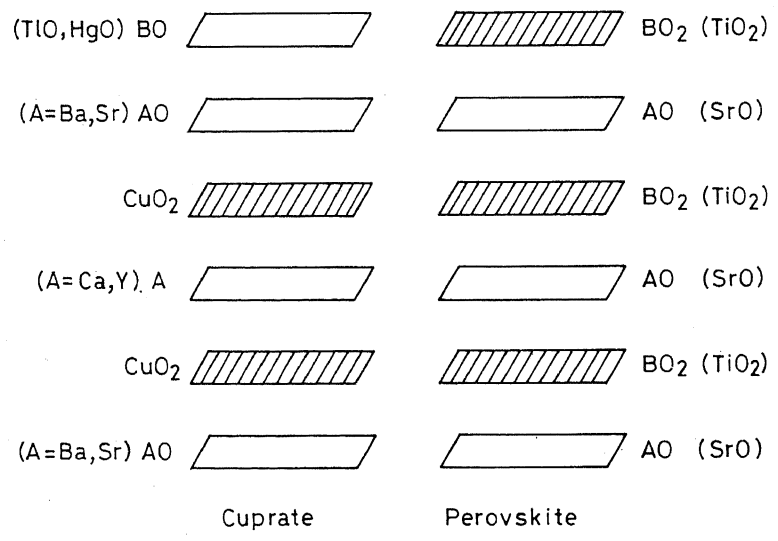


Figure 1. Layered lattice structure of perovskites and cuprate superconductors.

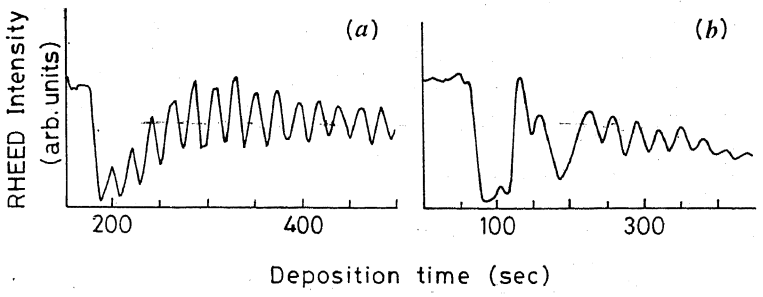


Figure 2. RHEED oscillations (0.4 nm) observed during the epitaxial growth of (*a*) $SrTiO_3$ (0 0 1) and (*b*) $BaCuO_2$ (0 0 1) on $SrTiO_3$ (0 0 1). Growth units are SrO/TiO_2, respectively (after Kawai *et al.*[5]).

possibility of growing novel or even unknown oxides layer by layer. It should be possible to fabricate a variety of hybrid devices and stacked unit cells or layers of perovskites which have good lattice matching with oxide substrates besides Si. For example, layered cuprate superconductors can be constructed or modified by depositing two-dimensional lattices (or layers) of different compositions chosen appropriately. In Figure 1, we show layered lattice structures related to perovskite oxides and high-T_c cuprates. The quality of such lattices is established by RHEED, RBS, STM, AFM and other techniques. A typical example of atomically controlled epitaxy of oxides is provided by $BaCuO_2$, which is a layered cuprate[5]. In Figure 2, RHEED oscillations obtained in the growth of such oxide layers are shown.

The wide range of properties of oxide systems certainly seems attractive compared to the Si/SiO_2 system, where the oxide can only provide it insulating characteristic. Valence control and impurity doping, readily accomplished in the case of Si, should be possible in oxides as well. In the short term, however, it should be possible to fabricate superconductor–insulator-superconductor (SIS) tunnel junctions exhibiting clear I–V hysteresis above 77 K. Such an SIS junction would be useful in switching devices which are expected to work faster than conventional transistors. It should be noted that there has already been success in fabricating simple devices using superconducting films or grain boundary and step edge junctions[6,7]. The use of oxides in combination with Si has also several possible applications in integrated devices. Clearly, there is much to be done in exploring oxides for electronic applications and in designing oxide superlattices with unexplored quantum functions. We have to find ways of doping, controlling valency and attaining the right carrier mobilities as well as in designing the right structures.

The importance of oxide electronics has been recognized widely in Japan, where there is considerable investment in this area. A large number of chemists, physicists and engineers are working on the various aspects of metal oxides, starting from synthesis and characterization to device applications. The Japanese consider oxides as electronic materials of the 21st century. In Tokyo Institute of Technology alone, a dozen or more professors work in this area. It is highly desirable that strong interdisciplinary groups in this country are encouraged and supported to work on various aspects of oxides.

1. Rao, C. N. R. and Raveau, B., *Transition Metal Oxides*, VCH Publishers, New York, 1995.
2. Rao, C. N. R., *Annu. Rev. Phys. Chem.*, 1989, **40**, 291.
3. Bednorz, J. G. and Müller, K. A., *Z. Phys.*, 1980, **B64**, 189.
4. Chanara, K., Ohno, J., Kasai, M. and Kozono, Y., *Appl. Phys. Lett.*, 1993, **63**, 1990; McCormack, M., Jin, S., Tiefel, T. H., Fleming, R. M., Phillips, J. M. and Ramesh, R., *Appl. Phys. Lett.*, 1994, **64**, 3045; Mahesh, R., Mahendiran, R., Raychaudhuri, A. K. and Rao, C. N. R., *J. Solid State Chem.*, 1995, **114**, 297; Mahendiran, R., Mahesh, R., Raychaudhuri, A. K. and Rao, C. N. R., *Pramana*, 1995, **44**, L393.
5. Kawai, M., Liu, Z., Sekine, R. and Koinuma, H., *Jpn. J. Appl. Phys.*, 1993, **32**, L1208; Gupta, A., Shaw, T. M., Chern, M. Y., Hussey, B. W., Guloy, A. M. and Scott, B. A., *J. Solid State Chem.*, 1995, **114**, 190; Shaw, T. M., Gupta, A., Chern, M. Y., Batson, P. E., Laibowitz, R. B. and Scott, B. A., *J. Mater. Res.*, 1994, **9**, 2566.
6. Dimos, D., Chaudhuri, P., Mannhart, J. and LeGoues, F. K., *Phys. Rev. Lett.*, 1988, **61**, 219.
7. Daly, K. P., Dozier, W. D., Burch, J. F., Coons, S. B., Hu, R., Platt, C. and Simon, R. W., *Appl. Phys. Lett.*, 1991, **58**, 543.

C. N. R. Rao is in the Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore 560 012, India.

Noncentrosymmetric monolayers from centrosymmetric molecules – A fresh approach to materials for frequency doubling

T. P. Radhakrishnan

The age we live in is often called the 'electronic age'. Some recent developments in physics, chemistry, materials science and engineering, however, indicate that a 'photonic era' is not too far in the future, when the role of electrons in information technology will be played much more efficiently by photons. Signs of the transition can already be seen, e.g. in fibre-optic communication systems utilizing electrooptic devices; many believe that all-photonic devices will soon appear. These devices for light transmission, processing and storage would be based on a variety of nonlinear optical (NLO) phenomena. The nonlinearity refers to the second- and higher-order terms in electric field (E) in the expres-

sion for the bulk electric polarization ($\boldsymbol{P}$) of a medium. $\chi^{(1)}$ refers to the linear electric susceptibility:

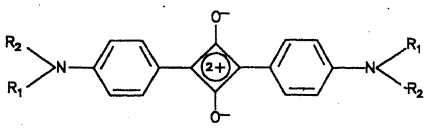
$$\boldsymbol{P} = \chi^{(1)}\boldsymbol{E} + \chi^{(2)}\boldsymbol{E}\cdot\boldsymbol{E} + \chi^{(3)}\boldsymbol{E}\cdot\boldsymbol{E}\cdot\boldsymbol{E} + \ldots$$

and the other $\chi^{(n)}$ produce the different NLO effects; for example, $\chi^{(2)}$ leads to phenomena such as frequency doubling, also called second-harmonic generation (SHG). In a centrosymmertric medium, reversal of $\boldsymbol{E}$ should cause exact reversal of $\boldsymbol{P}$, and, as seen from the above expression, this is possible only if all $\chi^{(n)}$ with even n are zero. In other words, nonzero even-order χ can be realized only in noncentrosymmetric systems. Frequency doublers in use today employ inorganic crystals like potassium dihydrogen phosphate, lithium niobate and β-barium borate which have noncentric arrangement of the atomic lattice. Recent efforts have been directed at exploiting the large hyperpolarizabilities of selected organic molecules (like donor–acceptor-substituted benzenes, quinones, stilbenes, etc.) gathered in a suitable noncentrosymmetric array in a crystal, polymer matrix or thin film. Large molecular hyperpolarizabilities which lead to high bulk $\chi^{(2)}$ values are usually associated with strong intramolecular charge transfer in the molecular dipoles.

Obtaining a noncentrosymmetric array of molecules is a challenging problem. A noncentrosymmetric molecule with a permanent ground state dipole moment does not guarantee a noncentrosymmetric crystal lattice. On the contrary, they often form centrosymmetric crystals. However, it is the common experience so far that if one is to fabricate noncentrosymmetric lattice, one should build it with a noncentrosymmetric molecule, not a centrosymmetric one, though there are rare exceptions where centrosymmetric molecules with octupolar moments have formed noncentric lattices which showed moderate SHG. Crystal engineering utilizing H-bond interactions, chirality, etc., to build noncentric lattices, growth of Langmuir–Blodgett (LB) films of the X or Z type, where molecules are deposited in a head-to-tail fashion and electric field poling of molecular dipoles embedded in a matrix or attached to polymer chains are some of the common approaches to obtain noncentric arrangement necessary for SHG materials.

With this background the recent report by Ashwell and coworkers[1] on the detection of very strong SHG from centrosymmetric squarine dye molecules (Figure 1) deposited as LB films comes as an interesting surprise. Crystal structure of **1** is found to be centrosymmetric, belonging to the $P2_1/c$ space group and the diphenyl-substituted squarine moiety is flat and centrosymmetric. As expected, the crystals showed no SHG but the LB films (though of poor quality) were found to be SHG-active. Ashwell and coworkers report that **2** and similar systems (with R = other alkyl groups) provide better films and reproducible SHG which is among the highest seen so far in LB films. Reflectivity studies indicated that the LB films consisted of monolayers with the chromophores lying nearly parallel to the substrate surface. Absorption spectra of the films showed characteristic bands due to intermolecular charge transfer in the 660–690 nm range depending on the film deposition pressure.

The crystals of **1** consist of noncentric chromophore dimers which arrange themselves in a centrosymmetric lattice. The authors, therefore, suggest that in the LB films of **2** and similar molecules, noncentric dimers may be present which, however, form a noncentrosymmetric array, thus enabling $X^{(2)}$ to be nonzero. Since the chromophores themselves are centrosymmetric, intramolecular charge transfers do not contribute to SHG. Interfacial effects cannot give rise to such strong SHG as seen in these LB films. Therefore, the authors attribute the strong SHG to intermolecular charge transfer interactions in the dimer units. These charge transfers occur between the dialkylaminophenyl group of one molecule and the squarine moiety of the other in the dimer unit. It is also notable that a good correlation exists between the absorption wavelength variation from 656



1 $R_1 = R_2$ = butyl

2 R_1 = methyl, R_2 = hexyl

Figure 1. Squarine dye molecules deposited in LB films by Ashwell *et al.*[1].

to 694 nm and the increase of SHG efficiency as a function of film deposition pressure. The compression process appears to influence strongly the orientation of the chromophores (or chromophore dimers) in the noncentrosymmetric array.

The LB film structure is not fully understood and, therefore, the mechanism by which the strong SHG arises has to be treated as tentative. Film characterizations were carried out on samples deposited on silver-coated slides, but different film characteristics for samples deposited on plane glass slides used in the SHG studies cannot be ruled out. Though some of these questions remain, the discovery of Ashwell and coworkers is notable for the following important insights gained. There is no reason to believe any more that centrosymmetric molecules cannot be coaxed into noncentric arrays useful for SHG. Intermolecular charge transfers can be an efficient mechanism to obtain large bulk optical nonlinearities. This work is likely to initiate a fresh approach to the fabrication of quadratic nonlinear optical materials. The technological promise of these materials is enticing enough to justify detailed investigation and appraisal of such novel ideas.

1. Ashwell, G. J., Jefferies, G., Hamilton, D. G., Lynch, D. E., Roberts, M. P. S., Bahra, G. S. and Brown, C. R., *Nature*, 1995, **375**, 385–387.

T. P. Radhakrishnan is in the School of Chemistry, University of Hyderabad, Hyderabad 500 046, India.

Targeting RNA in the cell

S. C. Lakhotia

The typical textbook description of a cell suggests that the genetic information residing in the nuclearly located DNA is transcribed in a regulated fashion and the resulting RNA is processed and transported to cytoplasm, where it is translated and the protein so made is translocated to specific parts of the cell for its structural and/or enzymatic role in the cell's phenotype. However, such simplistic descriptions do not reveal the highly ordered three-dimensional organization of a cell, where every molecule, in fact, has a defined dynamic location. While localized distribution of at least certain specific RNAs in cells had been known for some time, it has become possible only recently to ask how this localized distribution is actually achieved. Application of molecular cell biological approaches has permitted newer insights into the three-dimensional molecular architecture of a cell and how the various molecules are vectorially transported within it to achieve the polarities that a cell has.

It is now clearly established that not only is the primary act of transcription of a given DNA sequence regulated with a certain degree of fidelity, the processing and transport of the transcripts is also a highly regulated process. Studies during the past few years have shown that certain RNAs are targeted to specific subcompartments of the cell where their translational product/s are immediately needed. The importance of such highly specific subcellular localization of RNA and its protein product is most obvious in the zygote cell since the unfolding of the entire body organization during the development of the zygote depends on the molecular asymmetries that are established in the unfertilized or the newly fertilized egg. In what follows, we discuss some recent studies which are attractive for their use of novel combinations of microscopic and molecular techniques to examine (i) the structural components of a cell that transport RNA away from the site of synthesis in nucleus to its final location in a specific subcompartment of the cell, (ii) the signals that tell the cellular transport machinery where to deliver the given RNA molecule and (iii) the components that keep the RNA anchored to the desired subcompartment.

mRNA transport and localization in cells: visuals of the process and mechanism

It has been known that the β-actin protein as well as its mRNA colocalize at the leading edge of lamellipodia in a variety of cell types[1–3]. Apparently, the highly localized distribution of β-actin mRNA in these cell types provides for the compartmentalization of actin synthesis. It has been shown in several cases[4] that the RNA localization signal (zipcode) is present in the 3′ untranslated region (3′ UTR) of mRNA. Kislauskis *et al.*[5] identified the specific 'zipcode' sequence/s by fusing different lengths of the 3′ UTR of β-actin to the reporter β-galactosidase (*lacZ*) gene; this allowed a direct visualization of the intracellular distribution of β-galactosidase activity in transfected chicken embryo fibroblasts by simple X-gal (a chromogenic substrate for β-galactosidase) staining. It was shown that two sequence elements in the 3′ UTR of β-actin mRNA provided the peripheral localization signal to the heterologous β-galactosidase mRNA so that in cells transfected with such chimeric constructs, X-gal staining revealed blue colour at the same place where the β-actin protein was. Of these two peripheral localizing signals, a 54-nucleotide sequence was the main 'peripheral RNA zipcode' while the homologous but less active 43-nucleotide sequence, present further downstream, was composed of 'zipcode elements'. The basic elements of these 'zipcodes' were identified as AATGC and GGACT, which act in concert so that multiple copies of each provide a stronger peripheral localization signal. These authors suggest that the high A/C content in the 54-nucleotide zipcode may also have a role in peripheral localization of the β-actin mRNA. It was also seen that addition of complementary oligonucleotides against the RNA zipcode delocalized the endogenous β-actin mRNA and altered the organization and shape of lamellipodia and stress fibres, without affecting levels of the β-actin mRNA or protein; this showed that the localization of β-actin mRNA is essential for the phenotype of these cells.

An answer to how the mRNA molecules may possibly remain anchored in cell cytoplasm, essential for the localized distribution, has been provided in a very ingenious study by Bassell *et al.*[6]: single mRNA molecules were seen with an unprecedented resolution and accuracy to be anchored on actin filament intersections in fibroblasts. To achieve this feat, they used different sized gold particles to 'see' simultaneously, at the ultrastructure level, the labelled oligo-dT probe hybridized *in situ* with polyA tail carrying mRNAs and the antibody bound to actin fibres. Furthermore, the oligo-dT probe hybridized with the polyA tails of cytoplasmic mRNAs was used as primer for *in situ* reverse transcription; the *in situ* extended cDNA chains and the oligo-dT–polyA hybrids were covisualized by immunobinding of different sized gold particles. These results unambiguously demonstrated the localization of individual mRNA molecules specifically at intersections of actin filaments. These studies also revealed that the mRNA at these intersections remains in a circular rather than an extended or linear conformation; this may help explain the interactions between the 3′ and 5′ ends of mRNA that are important in regulating translation.

Directed transport and localized anchorage of mRNA during oogenesis in *Drosophila*

The stages of oogenesis provide a remarkable material for study of localized RNAs and the mechanisms that regulate the spatial order in the distribution of different transcripts in different regions of oocyte. It is this asymmetric but highly ordered distribution of different macromolecules in the egg that determines the subsequent cascades of determination events in early embryonic development. The large body of genetic data has made *Drosophila* eggs a material of choice for such studies.

During oogenesis in *Drosophila*, the growing oocyte is nourished and provided with all its RNAs and proteins by the 15 sister nurse cells and the surrounding follicle cells. The nurse cells synthesize the different RNAs and proteins which are transported through intercellular connections to the growing oocyte, where each of these gets localized in a highly ordered and reproducible manner. Among the several different RNAs that are localized in *Drosophila* oocyte and have important roles in embryonic development, those produced by the *bicoid*, *oskar* and *nanos* genes are critical for anterior–posterior axis determination[7]. The *bicoid* (*bcd*) RNA in mature oocyte is localized to the anterior margin and remains confined to the anterior region of embryo, where it directs the development of anterior structures[7,8]. Likewise, transcripts of the *nanos* and *oskar* (*osk*) genes are located at the posterior pole of the egg and are crucial for the posterior development[7].

The localized distribution of these nurse-cell-produced transcripts in oocyte results from a series of events during oogenesis; a number of earlier acting genes are known to affect movement and the characteristic localization of these polarly located early determinants or morphogens[7]. The *staufen* (*stau*) gene plays a critical role in the last step of *bcd* RNA localization, i.e. its release from the egg cortex of mature oocyte into anterior cytoplasm of activated egg and early embryos. *stau* is also important in the final stages of posterior localization of the *osk* mRNA in the egg[7].

Pokrywka and Stephenson[9] found that *bcd*, *osk* and certain other RNAs were dynamically associated with cytoskeletal elements during different stages of oogenesis. These authors developed a simple procedure for fractionation of detergent-insoluble cytoskeleton from the soluble components of developing oocytes and showed that the association of *bcd* RNA with detergent-insoluble pellet was dependent on the stage of oogenesis. In the mature oocyte when *bcd* transcripts remain in the egg cortex at the anterior end, all of these were bound to cytoskeleton elements, but after egg activation, most of the *bcd* RNA was recovered in the soluble fraction. Like the *bcd*, the *osk* transcripts also showed a dynamic association with cytoskeletal components during oogenesis.

Identification of the various proteins associated with the pellet and soluble fractions through Western blotting led Pokrywka and Stephenson[9] to conclude that microtubules, and not microfilaments, are required for the assorting of *bcd* RNA in the pellet fraction and the pellet fraction is enriched in components required for stable anchorage of *bcd* RNA rather than in elements required for initial localization events like transport or docking. It was also clear that the localization of *bcd* transcripts in nurse cells and in deeper cytoplasm of egg involves different mechanism than localization in the anterior cortex of mature oocyte, which alone was pellet-associated.

The *stau* protein has a very interesting role in localization of both *osk* and *bcd* RNAs at opposite ends of the egg. In the case of *osk* RNA, it has been known that the *stau* protein is associated with *osk* RNA to mediate its binding with cytoskeletal components that transport the complex to the posterior pole[10]. Other studies have shown that the plus ends of microtubules within the oocyte are polarized towards the posterior pole[11]. A visual demonstration of the posterior facing of the plus ends of microtubules in oocyte was provided by Clark *et al.*[12] in flies transformed with a chimeric gene having the kinesin domain on its 5′ end and the β-galactosidase coding region on the 3′ end: this fusion protein retains the properties of kinesin (moving along microtubules towards their plus ends) and also shows β-galactosidase activity, which can be cytochemically visualized as blue staining with the X-gal substrate. Since all the blue staining β-galactosidase activity was seen to move posteriorly (due to the kinesin domain) in oocytes of these flies, it was obvious that the microtubules in oocytes are highly polarized with their plus ends pointing posteriorly. Viewed in this context, it becomes clear that binding of the *stau* protein with *osk* transcripts facilitates attachment to microtubules for transport of the *osk* transcripts to the posterior pole. Once the *osk* transcripts are delivered at the posterior pole, the *stau* protein is freed and quickly recruited for *bcd* localization.

Ferrandon *et al.*[13] have very convincingly documented a mutual binding of the *stau* protein with the *bcd* RNA and the interaction of this complex with microtubules. These authors used an RNA injection assay to map the 3′ UTR of *bcd* RNA necessary for anterior localization and to visualize its specific interaction with the *stau* protein. *In vitro* transcribed *bcd* RNA (full-length mRNA or the 3′ UTR or variously mutated 3′ UTR) was microinjected at ectopic sites in 0 to 1 h-old embryos (the endogenous *bcd* transcripts at this stage are anteriorly localized), and following a brief chase, the embryos were fixed to immunologically monitor the distribution of *stau* protein: the normal *bcd* 3′ UTR rapidly recruited the *stau* protein around itself to result in microscopically visible aggregates which moved during cell cycle stages in microtubule-dependent manner. Using deletions and linker-scanning substitutions, three noncontiguous localization signals in stem III and distal regions of stems IV and V of the predicted secondary structure of the *bcd* 3′ UTR were identified. The *stau*-dependent localization of *bcd* RNA at the time of egg activation and its earlier *stau*-independent localization to anterior cortex of oocyte use different but overlapping sequence signals. Aggregation of the cellular *stau* protein in early embryos with the microinjected appropriate *bcd* 3′ UTR RNA and their movement with microtubules showed that the *bcd* transcripts and the *stau* protein are dependent on each other for complex formation and localization. It is interesting to note that during oogenesis the *bcd* transcripts and the *stau* protein do not recognize each other; at that time the *stau* protein is involved in localizing the *osk* transcripts. Only after delivering the *osk* transcripts at the posterior pole, does the *stau* protein quickly move to the anterior of egg and bind with the *bcd* RNA for its localization and anchorage at the anterior region. Since in both the oocyte and embryos the localization of the *stau* protein depends on the appropriate RNA (*osk* in oocyte and *bcd* in early embryos) and *vice versa*, Ferrandon *et al.*[13] suggest that perhaps the same microtubule motor is involved in both these transcript movements. However, the events that make the *stau* protein change its affinity from *osk* to *bcd* transcripts in the mature egg are not known. Also, if the *stau–bcd* complex has to move on the same microtubule motors as the *stau–osk* complex, a change in the microtubule polarity may be necessary.

The myelin basic protein mRNA when microinjected into mouse oligo-

dendrocytes also aggregates into large particles that are transported along microtubules in anterograde direction[14]. It is very likely that many more such examples will be known soon and a class of proteins will be identified that recognize the 'RNA zipcodes' and accordingly transport the RNA with the help of cytoskeletal components and deliver them at the specified destination. While microtubules may be principally involved in the transport, the actin filaments may have a greater role in anchoring the localized transcripts. Also, as in the case of *osk* and *bcd* transcript localizations during different stages of oogenesis, different subsets of cytoskeletal components may be involved in transport as well as anchorage of the same RNA in a stage-dependent manner. With the feasibility of doing sophisticated biochemistry and molecular probing *in situ* in the cell at light as well as electron microscopic levels, it should be possible in the near future to have full graphic details of the movement of specific transcripts in cells to their destination and their anchorage after delivery.

1. Lawrence, J. B. and Singer, R. H., *Cell*, 1986, **45**, 407–415.
2. Hoock, T. C., Newcomb, P. M. and Herman, I. M., *J. Cell Biol.*, 1991, **112**, 653–664.
3. Hill, M. A. and Gunning, P., *J. Cell Biol.*, 1993, **122**, 825–832.
4. Singer, R. H., *Curr. Biol.*, 1993, **3**, 719–721.
5. Kislauskis, E. H., Zhu, X. and Singer, R. H., *J. Cell Biol.*, 1994, **127**, 441–451.
6. Bassell, G. J., Powers, C. M., Taneja, K. L. and Singer, R. H., *J. Cell Biol.*, 1994, **126**, 863–876.
7. Ding, D. and Lipshitz, H. D., *Bioessays*, 1993, **15**, 651–658.
8. Driever, M., in *The Development of Drosophila melanogaster* (eds Bate, M. and Martinez-Arias, A.), Cold Spring Harbor Laboratory Press, New York, 1993, pp. 301–324.
9. Pokrywka, N. J. and Stephenson, E. C., *Develop. Biol.*, 1994, **166**, 210–219.
10. St Johnston, D., Brown, N. H., Gall, J. G. and Janisch, M., *Proc. Natl. Acad. Sci. USA*, 1992, **89**, 10979–10983.
11. Theurkuf, W. E., Smiley, S., Wong, M. L. and Alberts, B. M., *Development*, 1992, **115**, 923–936.
12. Clark, I., Giniger, E., Ruohola-Baker, H., Jan, L. Y. and Jan, Y. N.,*Curr. Biol.*, 1994, **4**, 289–300.
13. Ferrandon, D., Elphick, L., Nusslein-Volhard, C. and St Johnston, D., *Cell*, 1994, **79**, 1221–1232.
14. Ainger, K., Avossa, D., Morgan, F., Hill, S., Barry, C., Barbarese, E. and Carson, J., *J. Cell Biol.*, 1993, **123**, 431–441.

S. C. Lakhotia is in the Department of Zoology, Banaras Hindu University, Varanasi 221 005, India.

OPINION

Declining phase of quality research in our universities

Samir Bhattacharya

A ripple is taking the shape of a vigorous wave, we have ignored it for long and allowed it to grow, and were not careful enough about the power of its irreparable damage. I am talking about the declining phase of our university research environment. Everyone in the academic domain knows in what way the job of a university teacher differs from a college teacher. Teaching and research is obligatory to the university teacher while the latter is not so for a college teacher. To sustain and strengthen the academic environment in higher education, research activity is a must, as feedback from research will keep the standard of teaching high and the creativity of a teacher will encourage the students of higher education to develop their finer qualities. Although these are known, what we are witnessing in our country is the fast receding of such a notion in university education. Doubtlessly, it is the university to be blamed for this deterioration; more precisely, it is the university teachers. Everyone is pointing to the degeneration of research quality in our universities, but there is no serious effort as yet to recuperate the university from this situation. We have not so far been conscious enough about the fact that such a trend of low-quality research activity would destroy our university system, which is the primary productive centre of higher education. Masters degree is the essential passport for research career in any institution of this country.

Let us first see what could be the possible reasons for this situation. The present policy of merit promotion in the universities is a major factor, where quality is not the regulatory factor. A university teacher has to wait for a specific time period and her/his potentiality is then judged by the number of papers published and the number of Ph D's produced. Except for a few universities, in most of the universities the number of publications is one of the major criteria, not the quality of publications. This has resulted in the mushroom growth of private journals (published by a person), where peer review system does not operate and money is the prime objective. If one pays the publication charges, a paper will be immediately published. Such an easy way of increasing the number of publications and then claiming the promotion has two pronged injurious effects on the academic environment; (i) it substantially reduces the desire for quality research among the teachers and (ii) it dampens the spirit of a young incumbent in the university. One may presume it to be impossible as members of selection committee should definitely scrutinize the quality of publications. Unfortunately, this usually does not happen in the majority of universities even if the post is an open post. The system of selection sometimes includes allocation of marks for the number of publications and not for the quality of publication. A person who may have two high-quality publications in well-reputed international journals may not be selected against a person who has ten publications in unrefereed low-quality journals. Another harmful trend is the yield to local pressure by the university authorities where the presence of political milieu worsens the

situation. In such an atmosphere it is difficult to select a genuinely strong academic person with good research background; consequently, many a time the best person desiring to start a career in the university suffers from frustration. The cumulative result of all these is leading the universities towards a bad to worse situation. However, after entering the university as a lecturer or an assistant professor, a person finds that for promotion quality is not the criterion, quantity is the major thrust and therefore her/his attempt is to produce more papers in unrefereed, below-the standard journals and low-grade Ph D's. She or he will probably get more persons in support from the faculty than a person with the will of producing quality research. Thus, when more faculty members are on the other side, a real potential creative researcher does not find a congenial environment and of the two options, i.e. submitting to the other side or facing the odd environment throughout life, the former appears to be easier.

The total aspect has a snowball effect, the research students trained by these teachers will never look for quality research, they will rather waive the possibility of good research to dominate the situation. Recruitment is naturally being influenced by this faculty structure and by all chance one of the low-grade Ph D's will get the opportunity and such a person is expected to repeat the bad trend and this will go on. This is exactly what is happening today.

Another crucial drawback of the universities to create research environment is their poor infrastructural facilities. Universities are suffering from acute financial crisis. A major portion of their sanctioned money goes to salaries and wages, the remaining thin portion is not even enough to provide a reasonable infrastructure for teaching. Research assistance, therefore, is being kept at abeyance. At least a few university teachers who really want to do good research are not getting any support from the university. To them one of the criteria for the job appears to be meaningless: how can one pursue research in this age in such a poor infrastructural condition. Although they might have great potentialities, they are unable to fight this odd environment and ultimately submit to low-quality research to keep their promotion avenue open. In this way universities are losing capable and strong research workers, which is also a great loss to the country. One can argue that a person may develop infrastructural facilities through research projects from different national agencies. At the beginning of research career or without individual research contributions in a specific line, a grant to develop research facilities is rather difficult. The situation in the IITs and agricultural universities is somewhat different. One of my colleagues curtly replied to my above criticism by mentioning that after spending a huge grant on national laboratories, what is our return today. I have simply pointed out to him that one cannot improve his position by simply pointing out the fault of others.

However, what could be the steps to salvage universities from this degenerative stage. Everyone will feel the immediate necessity of such steps, not only to save our universities from the peril, but also to save the future of our nation. I have a few suggestions:

(1) It may be recalled that in the recent past the practice of the University Grants Commission to sanction university departmental posts and necessary financial assistance was to evaluate the academic progress (which included both teaching and research curricula) of the department by means of a visiting committee, followed by the recommendation of this committee. This practice has been stopped for long. It is high time to invoke this practice again as evaluation of academic performance will compel the departments to improve their quality. This will provide an incentive to research since the committee may recommend open posts on certain subjects as per the performance.

(2) In the open posts, weightage should be given to the capability of a person instead of time-bound experience.

(3) There should be a ban on giving credit for publication in unrefereed journals.

(4) Assessment of the number of Ph D's produced by a teacher should be based on the publications from a particular thesis in refereed journals.

(5) The scope of Ph D registration by a part-time student (without any research fellowship) should be very limited as this category is now a major source in increasing the number of Ph D's of a university teacher. Part time Ph D registered students are usually not very much inclined to do quality research, their objective being just to get a degree, and university teachers, on the other hand, get an easy community to increase the number of Ph D's produced by them.

(6) Both for recruitment and promotion, citation index of publications should be given prime importance; this will enable recruitment and promotion of the best person among the applicants. Moreover, it will waive the present trend of publication in unrefereed journals; these journals are not indexed.

(7) Steps should be taken to provide infrastructural facilities for research activities.

(8) There should be an effective interaction between the university teachers and the research scientists of the national laboratories funded by CSIR, DST, DBT, ICMR, ICAR, etc. This will be mutually beneficial for university teachers and research scientists of national laboratories so far as the creativity part of research is concerned. Moreover, university teachers will be able to utilize the research facilities of national laboratories and scientists from those institutes may also take part in teaching as this will not only benefit the students but also provide the opportunity to the scientists to have exchange with bright students of the universities. Such an interaction is not difficult in the metropolitan cities, while building of national laboratories nearer to the non-metropolitan university premises may cause enormous gain to both research and teaching.

There may be other suggestions, the supreme duty for all of us being not to stand as passive onlookers to the deterioration of research activity in our universities. It is already affecting the teaching adversely and will soon take the shape of an incurable chronic disease which will destroy the future of our higher-education system, the worst sufferers being our future young generations. A day will come when a Master's degree holder in science will speak in a different language which will not be communicable to the advanced arena of science, and higher degree from the university will completely lose its credibility. Still there is time and hope, and steps to resuscitate the universities should be initiated immediately.

Samir Bhattacharya is in the Department of Zoology, Visva-Bharati University, Santiniketan 731 235, India.

Implications of recent levelling observations for Tehri and other high dams in the Himalaya

Seismic hazards around the Tehri and other existing and proposed high dams in the Himalaya are a matter of concern to many people. The magnitude and dimensions of the problem appear to increase with every new set of geophysical and geological data gathered from the Himalaya. But the flexibility and readiness of the people involved to improve their designs for the dams transparently in the light of the evolving perceptions about seismic hazards[1–5] is not evident to us at least. Still, we buttress further in this article the evidence for an aspect of seismic hazards in the Himalaya.

A set of repeat geodetic levelling observations[6] obtained in Nepal (Figure 1 and 2 *a*) during the last two decades vividly reveal the ongoing deformation of the Himalayan upper crustal rocks in response to the relative convergence of Indian and Eurasian lithospheric plates. Laterally localized near-surface deformation is superimposed on broader-scale crustal strains (Figure 2). We provide here an explanation for some of the more localized elevation changes observed in Nepal and consider their seismic hazard implications.

The Nepal levelling data

Jackson and Bilham[6] have recently published the results of two successive sets of levelling observations across the width of Nepal along the main north south highway through Kathmandu (Figure 1). The period elapsed between the two sets varied between 2 and 13 years. Figure 2 *a* is a display of the observed annual rates of elevation changes along a line normal to the local trend of the Himalaya. Figure 2 *a* is created by projecting the benchmark positions along the winding highway onto a horizontal straight line oriented along N 10°. Jackson and Bilham[6] have provided an interpretation of the relatively longer wavelength features of the curve displayed in Figure 2 *a* by postulating a slip rate of 5 mm/yr on the detachment (Figure 2 *c*), an intracrustal thrust fault type of surface of separation between the Himalayan rocks above and the relatively northward-moving Indian shield material below[7].

Data from only the southern Nepal are shown in Figure 2 *a* for three reasons. Firstly, distance-dependent random errors in levelling observations grow northward from the southern end of the levelling line[6]. Secondly, as the levelling line approaches the main central thrust (MCT), the effect of plate convergence accommodated along a major ramp in the detachment becomes prominent in the form of a large-amplitude, long-wavelength signal in the elevation change observations[6]. This is not of interest in the present context. Thirdly, most of the river valley projects and high dams that can be envisaged to develop the hydroelectric and irrigation potential of these hills are located in the Lesser and Outer Himalaya.

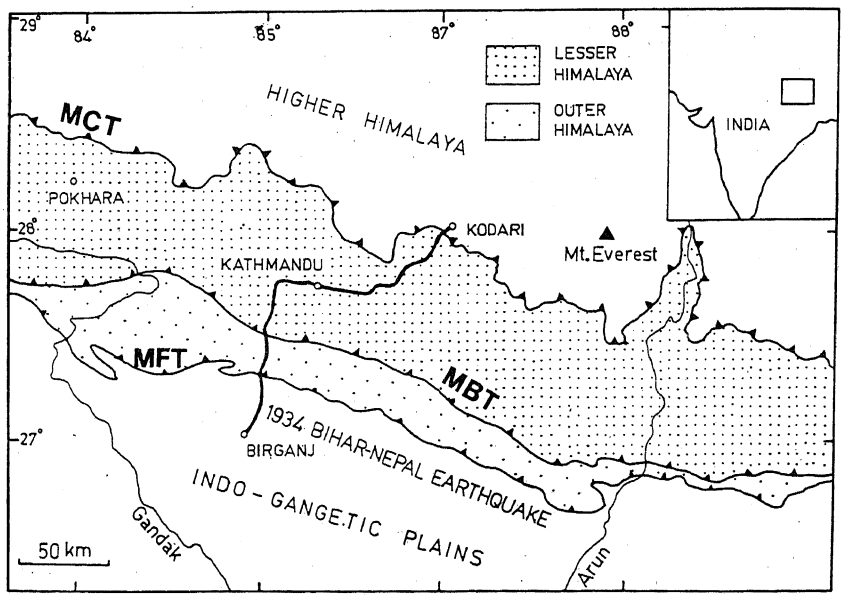


Figure 1. Location of the levelling line in Nepal. MFT – main frontal thrust; MBT – main boundary thrust; MCT – main central thrust.

Imbricate thrusts in the Nepal Himalaya

Several shorter-wavelength signals in the Nepalese levelling data deserve explanation also. We show in Figure 2 *c* our interpretation of the two most prominent of these features on the assumption that they could be associated with slip on relatively steeper near-surface thrust faults. Similar faults could be conceived for other less prominent short-wavelength features in the data. Such shallow-level and relatively steep dipping thrust faults are called imbricate thrust[8]. They occur in suites and are associated with a major, deeper thrust fault dipping at a comparatively gentler angle[8]. Although geological mapping in different parts of the Himalaya is still far from complete, numerous imbricate thrust faults have been mapped already[9, 10]. Since precise locations of benchmarks used by Jackson and Bilham[6] have been masked, it is not possible to associate the imbricate thrusts postulated by us in Figure 2 *c* with specific thrust faults mapped in the Nepal Himalaya. The possibility also exists that the geodetic observations may have revealed imbricate thrusts which are as yet 'blind' or concealed below the ground surface.

Role of imbricate thrusts during a great thrust earthquake

Since the observations reported in Figure 2 *a* were taken during a period when no major earthquake had occurred in the region, they reveal aseismic crustal defor-

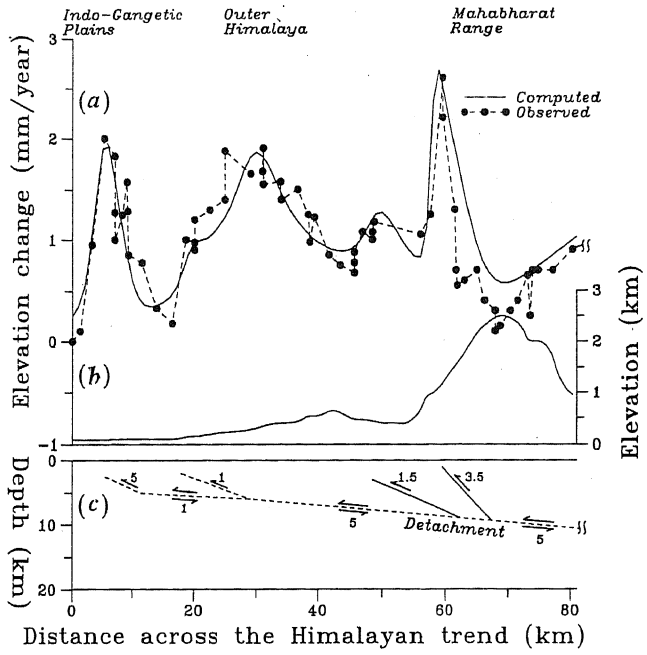


Figure 2. ***a***, Comparison of observed and calculated rates of ground elevation changes in southern Nepal. ***b***, Profile of topography along the projected line. ***c***, The model used to compute the results shown in (***a***). The dotted parts of the fault model were also considered by Jackson and Bilham[6] to explain the long-wavelength features. Solid lines indicate the imbricate faults suggested by us to explain two short-wavelength features in the observations. The numerals shown against faults represent the postulated annual slip rates (mm/yr) on the respective faults to explain the observations.

mation. It is not unreasonable to surmise that these imbricate thrusts will be reactivated also in the next great earthquake of the region. These faults could be loci of large relative slips between rocks across them. Significant ground level changes may occur in regions near their updip terminations. Concomitant high particle velocities and accelerations may be expected in these pockets during great earthquakes. Ample evidence of this has been documented in the literature.

Oldham[11] surveyed the changes in the ground surface brought about in the Shillong plateau region during the great earthquake of 12 June 1897. He reported surface-breaking faults near the northern border of the Shillong plateau. He estimated relative ground level changes of up to about 12 m near these faults. Although Oldham[11] estimated that the faults were subvertical, Gahalaut and Chander[12] have surmised that they could be imbricate to the major mid-crustal thrust fault in which the main rupture of this earthquake occurred. The 1897 earthquake caused the ponding of south-flowing rivers in the southern part of the Shillong plateau[11]. This too could be explained with coseismic activity on imbricate thrust faults under southern Shillong plateau[12]. Finally, Oldham[11] argued lucidly from the evidence of boulders thrown into air during the earthquake that ground acceleration exceeding g (acceleration due to gravity) occurred during that earthquake. We suggest that slip on a nearby imbricate thrust was involved in this case also. It is important to recall that Oldham[11] has given exactly the same explanation for these observations. However, he used the phrase 'secondary and tertiary reverse faults' instead of 'imbricate thrust faults'. The hazard implication of this observation of high acceleration has been remarked by Khattri[2] also.

Byrne *et al.*[13] investigated the source of the 1945 Makran earthquake and suggested that prominent uplift along a narrow coastal belt could be due to slip on a shallow fault imbricate to the deeper detachment-type plate boundary in the Makran subduction zone. Barrientos and Ward[14] and Plafker[15] also referred to significant activity on imbricate thrusts extending up from the main plate boundary thrust faults associated with the great Chile earthquake of 1960 and the great Alaska earthquake of 1964.

Implication of imbricate thrusts for civil engineering structures in the Himalaya – the concept of the nearest active fault

An important consideration for the design of a critical civil engineering structure, such as a dam or a nuclear reactor, especially from the standpoint of seismic hazards, is the distance of the site from the nearest active fault. Ordinarily, the distance between an engineering site and a surface-breaking active fault is determined on a geological map. However geodetic and other geophysical techniques, such as earthquake observations, can lead to identification of subsurface active faults also. Geodetic techniques have the advantage that even aseismically active faults can be detected.

Recent investigations of great Himalayan earthquakes have led to the consensus that they occur by extended ruptures in the detachment[7, 12, 16–20]. The depth of the detachment has been estimated to be less than 5 km under the main frontal thrust (MFT) and about 17 km under the MCT[7, 16, 18, 19, 21–23]. Thus, there is at least one active fault beneath every engineering site in the Outer and Lesser Himalaya. If some faults imbricate to the detachment could be active in this Himalayan domain, as has been remarked elsewhere on geological grounds[1, 10, 24–26] and shown to be the case in the Nepal Himalaya through geodetic observations, then they could be potentially even nearer to an engineering site than the detachment.

On the Tehri dam site

The Kumaun–Garhwal Himalaya, including the region around Tehri dam site in Garhwal (Figure 3), has been investigated geologically, especially in the recent decades. The monograph by Valdiya[10] is an important source book for the geology of the Lesser Himalaya of the Kumaun–Garhwal region. Valdiya[1, 10, 24–29] and others[30–32] have documented geomorphic and other geological evidence for neotectonic and even continuing tectonic deformation in different longitudinal and transverse segments of the Kumaun–Garhwal Himalaya. This includes suggestions about possible deformation along imbricate thrusts. See the discussion section below for a limitation of the geologi-

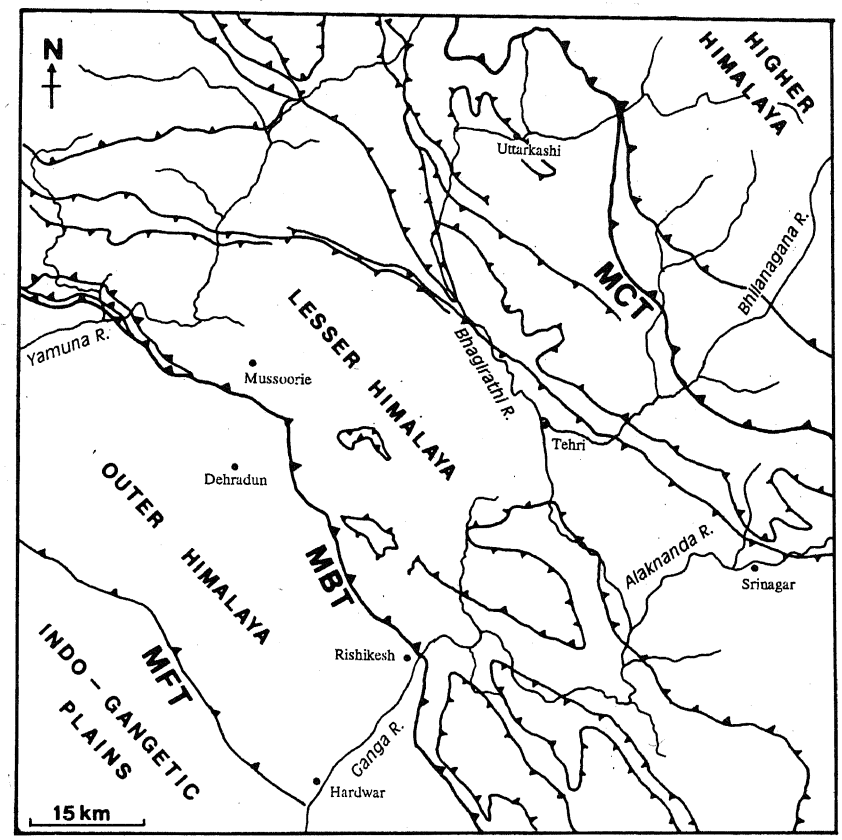


Figure 3. A randomly selected geological map of a section of the Lesser Himalaya showing some of the mapped thrust faults (after Valdiya[10]). Smaller unnamed faults hading northerly are imbricate faults.

cal inference regarding the activity of faults.

The 1991 Uttarkashi earthquake[33] and recent seismicity observations[34–35] provide further confirmation about the active or ongoing tectonics of the Garhwal region. There was damage at Tehri during the great Kangra earthquake of 1905 (ref. 36). The rupture models proposed by Seeber and Armbruster[7], Chander[19] and Gahalaut and Chander[16] suggest that the detachment under the NW Garhwal Lesser and Outer Himalaya slipped during the earthquake. Epicentres of many small and microearthquakes have been located within about 10 km of the Tehri dam site through observations made[34, 35] in the 1980s. Composite fault plane solutions for some of these earthquakes suggest activity along high-angle reverse faults within the upper 15 km of the crust[34, 35]. These seismologically inferred reverse faults too could be properly called imbricate thrusts. However, in the absence of sufficiently detailed geological mapping as well as the inevitable uncertainties in the estimates of hypocentral coordinates it is not possible to associate these seismologically inferred imbricate thrusts with specific geologically mapped faults of the region. Still in our opinion this is an evidence showing that at least some imbricate thrusts are active today close to the Tehri dam site.

Detailed designs for Tehri dam and appurtenant structures have been made already. Still we suggest that, in view of the magnitude and importance of the project, detailed repeat geodetic observations should be undertaken around the Tehri dam site to supplement the geological and seismological observations and to detect if in addition any buried or exposed faults near this site also are currently creeping aseismically as in the Nepal Himalaya. If such observations have been taken already by governmental agencies, then the results should be published so that interested scientists outside these agencies, may also assess them. The published levelling observations for the Saharanpur–Dehradun–Mussoorie highway[36, 37] are not relevant for this purpose because of the distance between the site and the closest benchmark.

Discussion

Bomford[38] in his well-known book on geodesy specifically recognized that geodetic surveys are useful at civil engineering sites to mark points on the ground with great accuracy. Bomford[38] also underlined the use of repeat geodetic observations for crustal deformation studies. Our analysis above illustrates the application of geodetic observations for identifying the locales of ongoing crustal deformation in the Himalayan setting specifically. Combining the results of sufficiently detailed geodetic observations with the geological and seismological observations should help in settling the question of the nearest active faults for various dam sites with greater confidence.

The main advantage of geodetic methods is that we can detect activity which has occurred during the period between successive occupations of the benchmarks. The geological procedures, on the other hand, indicate activity without being very specific about their dates except when geochronological evidence regarding the activity is forthcoming. Geodetic observations should complement seismicity observations by identifying currently active faults.

We recall here that geodetic observations have been collected in India for nearly two centuries[39]. But only a very small subset of these are repeat observations with a possible bearing on crustal movements in respective regions. There are three prime requirements when geodetic observations are to be used in this way. Firstly, the benchmarks should be closely spaced. Secondly, the reference benchmark should be well outside the region of suspected crustal deformation. Thirdly, the reoccupation of benchmarks should be at sufficiently short intervals of time. To the best of our knowledge there are no repeat geodetic observations in India which fit these requirements entirely. Repeat geodetic observations in connection with the 1897 Shillong plateau[11] and the 1934 Bihar–Nepal[40] earthquakes suffer from the fact that they were within or close to the epicentral tracts. Repeat levelling along the Saharanpur–Dehradun–Mussoorie line has revealed[36, 37] changes during and sub-

sequent to the Kangra earthquake of 1905. But the large distances between benchmarks, the long time gaps between successive observations and the desirability of shifting the reference benchmark at Saharanpur to a bed rock site further south have been remarked[41]. Finally, repeat geodetic observations to measure aseismic fault movements in the Hardwar[42], Kangra[43] and Mandi[44] regions had too limited a spatial extent in our opinion.

An important reason why geodetic methods have not been employed widely for dam site investigations is that in the past the procedures were slow and relatively costly. However, with the availability of the global positioning system these factors should no longer be disadvantages or deterrents for carrying out geodetic surveys as envisaged.

Conclusion

Our interpretation of recent levelling observations in the Nepal Himalaya suggests that a part of the current interseismic upper crustal deformation may be localized along shallow imbricate thrust faults. We have emphasized three implications of these geodetic observations and analyses. Firstly, similar imbricate faults may be slipping aseismically elsewhere in the Himalaya also. Secondly, during great earthquakes of the respective regions, coseismic slip along such faults may lead to locally strong ground motion with attendant large displacements and high accelerations. This could be a threat to nearby civil engineering structures, such as dams, as well as smaller, non-engineered constructions, such as individual houses in villages. Thirdly, there is need to expand repeat geodetic observations in all Indian segments of the Himalaya.

1. Valdiya, K. S., in *Earthquake Hazard and Large Dams in the Himalaya* (ed. Gaur, V. K.), INTACH, New Delhi, 1993, pp. 1–34.
2. Khattri, K. N., in *Earthquake Hazard and Large Dams in the Himalaya* (ed. Gaur, V. K.), INTACH, New Delhi, 1993, pp. 35–62.
3. Gaur, V. K., in *Earthquake Hazard and Large Dams in the Himalaya* (ed. Gaur, V. K.), INTACH, New Delhi, 1993, pp. 63–74.
4. Bolt, B. A., in *Earthquake Hazard and Large Dams in the Himalaya* (ed. Gaur, V. K.), INTACH, New Delhi, 1993, pp. 75–92.
5. Brune, J., *Tectonophysics*, 1993, **218**, 281–286.
6. Jackson, M. and Bilham, R., *J. Geophys. Res.*, 1994, **99**, 13897–13912.
7. Seeber, L. and Armbruster, J. G., in *Earthquake Prediction – An International Review*, American Geophysical Union, Washington, DC, 1981, vol. 4, pp. 259–277.
8. Price, N. J. and Cosgrove, J. W., *Analysis of Geological Structures*, Cambridge University Press, New York, 1990, p. 502.
9. Gansser, A., *Geology of the Himalaya*, Interscience, New York, 1964, p. 289.
10. Valdiya, K. S., *Geology of Kumaun Lesser Himalaya*, Wadia Institute of Himalayan Geology, Dehradun, India, 1980, p. 291.
11. Oldham, R. D., *Mem. Geol. Surv. India*, 1899, **29**, 1–379.
12. Gahalaut, V. K. and Chander, R., *Tectonophysics*, 1992, **204**, 163–174.
13. Byrne, D. E., Sykes, L. R. and Davis, D. M., *J. Geophys. Res.*, 1992, **97**, 449–478.
14. Barrientos, S. E. and Ward, S. N., *Geophys. J. Int.*, 1990, **103**, 589–598.
15. Plafker, G., *J. Geophys. Res.*, 1972, **77**, 901–925.
16. Gahalaut, V. K. and Chander, R., *J. Geol. Soc. India*, 1992, **39**, 61–68.
17. Yeats, R. S. and Lillie, R. J., *J. Struct. Geol.*, 1991, **13**, 215–225.
18. Ni, J. and Barazangi, M., *J. Geophys. Res.*, 1984, **89**, 1147–1163.
19. Chander, R., *Tectonophysics*, 1988, **149**, 289–298.
20. Chander, R., *Tectonophysics*, 1989, **170**, 115–123.
21. Lyon-Caen, H. and Molnar, P., *J. Geophys. Res.*, 1983, **88**, 8171–8191.
22. Chander, R., *J. Struct. Geol.* 1992, **14**, 621–623.
23. Gahalaut, V. K., Gupta, P. K., Chander, R. and Gaur, V. K., *Proc. Indian Acad. Sci. (Earth Planet. Sci.)*, 1994, **103**, 401–411.
24. Valdiya, K. S., in *Zagros–Hindukush–Himalaya: Geodynamic Evolution* (eds Delany, F. M. and Gupta, H. K.), American Geophysical, Union, Washington DC, 1981, pp. 87–110.
25. Valdiya, K. S., *Geology and Natural Environment of Nainital Hills, Kumaun Himalaya*, Gynodaya Prakashan, Nainital, 1988, p. 155.
26. Valdiya, K. S., *Curr. Sci.*, 1993, **64**, 873–884.
27. Valdiya, K. S., *Curr. Sci.*, 1991, **61**, 801–803.
28. Valdiya, K. S. *Curr. Sci.*, 1992, **63**, 289–296.
29. Valdiya, K. S., *Curr. Sci.*, 1994, **67**, 313–323.
30. Prasad, C. and Rawat, G. S., Proceedings of the International Symposium on Neotectonics in South Asia, Survey of India, Dehradun, 1986, pp. 301–312.
31. Khan, A. A., Dubey, V. S., Sehgal, M. N. and Awasthi, S. C., *J. Geol. Soc. India*, 1982, **23**, 392–401.
32. Nawani, P. C., Sanwal, R. and Khanduri, H. C., *J. Eng. Geol.*, 1991, **20**, 98–107.
33. Anonymous, *J. Geol. Soc. India*, 1992, **39**, 83–88.
34. Khattri, K. N., Chander, R., Gaur, V. K., Sarkar, I. and Kumar, S., *Proc. Indian Acad. Sci. (Earth Planet. Sci.)*, 1989, **98**, 91–109.
35. Kumar, Sushil, Ph D dissertation, Department of Earth Sciences, Univ. of Roorkee, Roorkee, 1994, pp. 231.
36. Middlemiss, C. S., *Mem. Geol. Surv. India*, 1910, **37**, 1–409.
37. Rajal, B. S., Virdi, N. S. and Hasija, N. L., Proceedings of the International Symposium on Neotectonics in South Asia, Survey of India, Dehradun, 1986, pp. 146–159.
38. Bomford, G., *Geodesy*, Clarendon Press, Oxford, p. 731.
39. Phillimore, R. H., *Survey of India*, Dehradun, 1945, vol. 1, pp. 1–400.
40. Dunn, J. A., Auden, J. B. and Ghosh, A. M. N., *Mem. Geol. Surv. India*, 1939, **73**, 27–48.
41. Chander, R. and Gahalaut, V. K., *Curr. Sci.*, 1994, **67**, 531–534.
42. Arur, M. G. and Hasija, N. L., Proceedings of the International Symposium on Neotectonics in South Asia, Survey of India, Dehradun, 1986, pp. 330–344.
43. Chugh, R. S., International Symposium on Recent Crustal Movement, Zurich, 1974.
44. Arur, M. G. and Singh, A. N., Proceedings of the International Symposium on Neotectonics in South Asia, Survey of India, Dehradun, 1986, pp. 353–365.

ACKNOWLEDGEMENT. Financial support of UGC (E&T) to V. K. Gahalaut is acknowledged.

R. Chander and V. K. Gahalaut are in the Department of Earth Sciences, University of Roorkee, Roorkee 247 667, India.

Storing river runoff in the Gulf of Khambhat: A critique of the 'Kalpasar' concept

A scheme of dam construction across the Gulf of Khambhat (Cambay) for storage of runoff water of several rivers draining into the Gulf has been proposed by Kane[1]. The proposed dam will connect Gundi near Bhavnagar on the west (Saurashtra side) and Hansot near Bharuch on the east (South Gujarat side) and is expected to create a 2900 km^2 freshwater lake. The total length of the dam is estimated to be 56.4 km and the cost approximately Rs 3000–4000 crores (~US$ 1000 million). The proposed lake has been named 'Kalpasar' (taken from a Sanskrit word meaning a lake which can give one whatever one wishes standing on its bank). The annual river inflow into the Kalpasar lake has been estimated to be about 14,000 million m^3 after the proposed Narmada dam and other irrigation schemes on the rivers Sabarmati, Mahi and other minor rivers have been completed.

M/s Haskoning, a Dutch firm, was retained by the Government of Gujarat and a reconnaissance report on this project was submitted in 1989. The proposal of Kane[1] is based on the report of Haskoning[2]. The scheme has also been advocated by Patel[3] as part of a larger strategy to develop water resources in the State of Gujarat.

While we generally endorse the concept of creating a freshwater lake in the Gulf of Khambhat, we suggest an innovative modification of the dam structure that may help to store more fresh water and additionally lead to considerable saving in the cost through reduction in both length and thickness of the dam structure. The proposed modification is based on the basic hydrological principle, namely, the significant density difference between the sea water and the river runoff, that should allow the lighter river water (density ~1.0 g/cc) to float on the denser sea water (density 1.025 g/cc). Employing this principle, if the dam structure is so designed that it is open from the bottom allowing hydraulic continuity on north and south of the dam, one should expect the incoming lighter river water to push the denser sea water out to the open ocean through the open bottom of the dam. The modification may also obviate some of the difficulties that we anticipate in the scheme as described by Kane[1], without affecting the possible benefits.

In the scheme proposed by Kane[1] an earth fill dam is to be constructed across the Gulf along a suitably selected contour so that the maximum depth of the water column does not exceed 30 m. At the time of the lowest tide level, sea water from the northern (landward) portion of the dam is to be drained out to the southern (seaward) side, to the extent possible through sluice gates. Subsequently, when the fresh water in the monsoon starts flowing into the northern portion through the rivers, this portion is expected to be converted into a freshwater lake through mixing and dilution of the remaining sea water with the in-flowing river water. In this manner after a period of a few (3–5) years the reservoir is expected to be washed off of all the residual sea water. In this scheme, the maximum water that can be stored in the Kalpasar without flooding any additional coastal areas is between the highest and the lowest tide levels. With an average of 14,000 million m^3 of water per annum, estimated to be available in the 2900 km^2 area of the lake, the maximum estimated thickness of freshwater column at the dam site is expected to be in excess of 6 m, which is of the same order as the tidal range (about 8 m) in the Gulf of Khambhat. Thus, even though a very large volume of reservoir will be created (up to 30 m depth at the dam site), the volume actually available for storage of fresh water would be smaller than the estimated freshwater inflow in several years of above-average rainfall. Therefore, the excess river water runoff would have to be either stored on land above the high tide level or would have to be let out into the open sea.

There is also some concern regarding mixing and dilution being able to wash off the residual sea water in 3–5 years time as envisaged by Haskoning[2] and Kane[1]. Simple calculations assuming the water body to be wedge-shaped between the northernmost section of the Gulf (~100 km from the dam) and the dam site, where the depth will be ~30 m, give a residence time of water in the lake (ratio of the total storage volume to the annual average inflow) of about 5 years. But for salt concentration to reduce from the sea water value of ~35,000 ppm to ~500 ppm (suitable for irrigation and drinking), the time required will be much longer even if perfect mixing and no evaporation (about 2 m water column) are assumed. Mixing and dilution, therefore, may not result in Kalpasar becoming a freshwater lake in a reasonable period after construction of the dam.

However, if we could construct the dam in such a way that the bottom 10 m or more is open over a fairly large length of the dam, say where the water column is more than 20 m, we may be able to obviate the above difficulty without compromising in any way the expected benefits. With the bottom of the dam open, there is no need to have gates to empty the reservoir. Also there is no need for having any spillway for draining off the excess river inflow out of the reservoir into the sea. Both these functions, namely (i) the draining of the sea water and (ii) any excess river inflow, will automatically be accomplished through the lower open end of the dam wall due to the transfer of the underlying sea water, governed by the hydrostatic balance on the two sides across the dam. The maximum height of the dam above the mean sea level in our scheme needs to be just enough to prevent the ingress of sea water overflowing the dam resulting from tide/wave-induced changes in the sea level.

The density contrast between the river runoff and the sea water is envisaged to maintain a reasonably well-defined interface at a depth of about 6 m or more to enable exploitation of the stored fresh water in a meaningful manner. This interface is, however, not static but moves up and down in response to hydrostatic changes induced (i) on the seaward side by the tides and waves and (ii) on the reservoir side by the inflow/consumptive use of the runoff water. Therefore, it may not be unreasonable to apprehend that due to mixing the interface may not be well defined, thereby defeating the very purpose of creating the reservoir at a huge expense. This is an aspect which

needs careful investigation employing both mathematical and physical models. However, a positive aspect of our scheme is that all inflows of fresh water into the reservoir not consumptively utilized will cumulatively depress the interface on the reservoir side and increase freshwater storage. Such inputs of fresh water to the Kalpasar lake also include rain (about 0.7 m) falling directly on to the lake and inflows from the minor rivulets, streams, creeks and channels and ground water discharge not accounted for in the 14,000 million m^3 of runoff from major rivers after the proposed dams on the Narmada and other rivers have been completed. We, therefore, intuitively anticipate that this effect involving cumulative lowering of the interface (with no spillage) will more than compensate for the possible mixing of saline and fresh water due to vertical movement of the interface induced by tides/waves. We expect that pressure changes due to the tides and waves can be expected to be communicated from the open sea side to the reservoir side, through the open bottom, rather gently in a piston-flow-like manner due to the damping of the disturbance with depth in the 20 m or more water column on either side of the dam. The dam itself will prevent the transfer of the wave/tide-induced turbulence at the surface from the open sea to the reservoir side.

The other factors influencing the position and thickness as well as movement of the fresh-water–sea-water interface in the reservoir, such as evaporation, consumptive use of water, wind-induced waves in the reservoir itself and the tidal effects within the body of the water stored in the reservoir, are common to both the schemes and, therefore, are not being discussed here. In our view, mixing is an area of major concern that needs a detailed investigation through mathematical models of currents and mixing processes likely to operate in the Kalpasar lake and the extent to which they may affect adversely the quality of water stored in the lake.

1. Kane, A. S., 'Kalpasar', the only permanent solution of the perennial water problem of Gujarat, Paper presented at the seminar on The Water Problems of Gujarat: Approaches to Solution, organized by the Centre for Research and Training in Rural Development (CERITA) and Gujarat Science Academy (GSA), 7 January 1995, PRL, Ahmedabad.
2. Haskoning, Khambhat gulf development reconnaissance report, Govt. of Gujarat, Water Resources Department, 1989.
3. Patel, V. J., Total water resources management in Gujarat for those who care for balanced growth, Paper presented at the seminar on The Water Problems of Gujarat: Approaches to Solution, organized by the Centre for Research and Training in Rural Development (CERITA) and Gujarat Science Academy (GSA), 7 January 1995, PRL, Ahmedabad.

S. K. GUPTA
P. SHARMA

Physical Research Laboratory,
Navrangpura,
Ahmedabad 380 009, India.
Also with Water Resources
Research Foundation,
c/o PRL,
Ahmedabad 380 009, India.

In situ preserved *Vertebraria* axes in Ib River Coalfield, Orissa

The Permian Gondwana form genus *Vertebraria Royle*[1] occurring from Post-glacial time through Late Permian is characterized by elongate, flattened, segmented, cylindrical casts which are simple or branched, often bearing root-like organs at the transverse ridges. In transverse section the *Vertebraria* axes show wedge-like sectors radiating from the centre[2]. According to Mussa[3] *Vertebraria* is a complex form consisting of several genera. Specimens of *Vertebraria* are found preserved in positions both parallel and perpendicular to the bedding plane and the actual nature of these axes was quite controversial for some time[3–9]. Most of the specimens assigned to *Vertebraria* are generally impressions or compressions and very rarely petrified specimens with anatomical details. It consists of a central region of exarch primary xylem with four to seven radiating arms of secondary xylem. The distinct growth rings with septate arrangement is surrounded by a cylindrical periderm with a well-developed cork, leaving spaces between the xylem arms. In longitudinal section the secondary xylem arms are connected at varying intervals by transverse platforms, also composed of secondary xylem tracheids, but generally containing a root trace.

Several specimens of *Vertebraria indica* Royle were collected from the Barakar beds exposed in a nala near Ganga Nagar village near Brijraj Nagar town, District Sambalpur, Orissa. No other element of Glossopteris flora is preserved in this bed. Some 50 specimens were found preserved as casts in the fine-grained, light pink, hard shales in an area of 2–3 m. The surface view showed 5–6 wedges 0.5–2.0 cm in diameter, with very little pith in the centre. The maximum length of these axes was 8.4 cm, lying diagonally along the bedding plane. Growth directions of these axes in the sediment indicate that *Vertebraria* was an underground root system of some plant. The actual plant which bore these underground axes was perhaps drifted to another site. Diagonal and vertical axes (Figure 1 *a, b*), preserved in a small area of 2–3 m, probably belonged to a single plant and represent *in situ* preservation. Total absence of horizontal axes and any other plant parts led us to believe this. The presence of such axes in various coal fields[10–14] indicates that some of the species of *Glossopteris* had arborescent habit supported by a well-developed strong root system. Wherever such vertical and diagonal axes were found, other plant organs like leaves were not found in nearby areas. Though horizontal axes along with other plant parts are commonly found in most coalfields, indicating autochthonous nature of coal, in no way are such vertical and diagonal *Vertebraria* axes indicative of autochthonous nature of coal as believed by some. Enormous quantity of vegetal matter which might

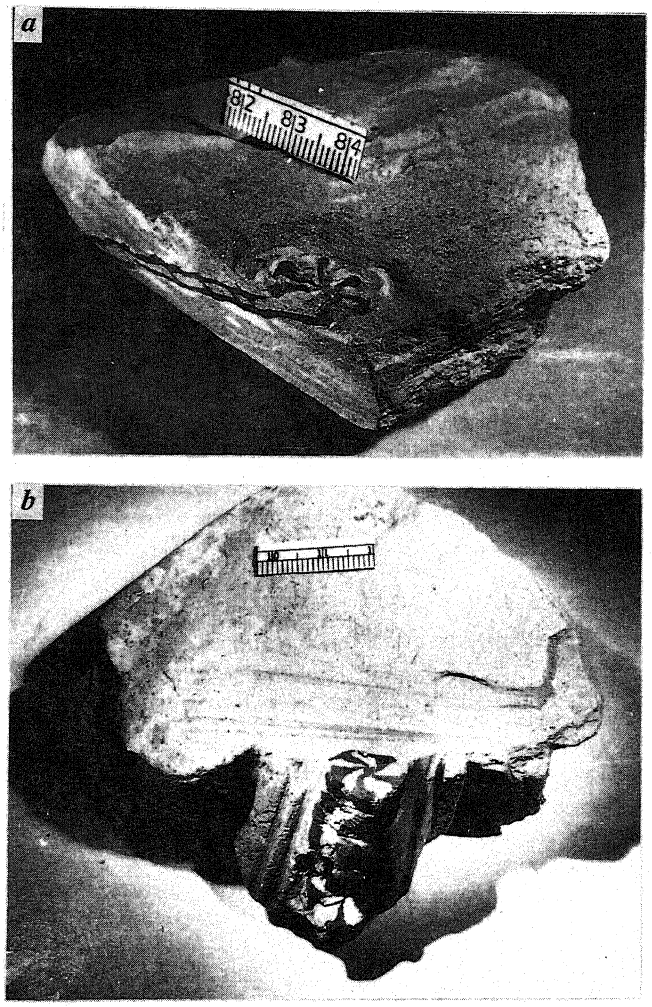


Figure 1 *a* **and** *b*. *Vertebraria indica* Royle, longitudinal sections of the specimens to show diagonally and vertically penetrating axes in the sediments, respectively. B.S.I.P. Museum specimen numbers 36863 & 36864.

have been required for peat formation of Gondwana coals cannot be sufficiently provided by few *in situ* preserved plants and their root system. It is logical to think that root system already embedded in the palaeosols had better chances of getting preserved *in situ* than other plant parts as found in Ib river and other coalfields of Lower Gondwana of India.

1. Royle, J. F., *Illustrations of the Botany and other Branches of Natural History of the Himalayan Mountains*, Wm. H. Allen and Company, London, 1833–39, vol. 2, p. 100.
2. Arber, E. A. N., *Catalogue of the Fossil Plants of the Glossopteris Flora in the British Museum*, Natural History, London, 1905, p. 255.
3. Mussa, D., *Boletim IG*, Instituto de Geociencias, 1978, **9**, 153–201.
4. Walkom, A. B., *Publ. Geol. Surv. Qld.*, 1922, **270**, 1–64.
5. Plumstead, E. P., Trans-Antarctic Expedition (London), 1962, Rept. 9, *Geology*, Pt. 2, p. 154.
6. Pant, D. D., *Phytomorphology*, 1968, **17**, 351–359.
7. Pant, D. D. and Singh, R. S., *Palaeontology*, 1968, **11**, 643–653.
8. Schopf, J. M., *Antarctic Res. Ser.*, 1965, **6**, 217–228.
9. Gould, R. E., *Gondwana Geology*, Austral. Nat. Univ. Press, Canberra, 1975, pp. 109–115.
10. Niyogi, D., *J. Sed. Petrol.*, 1966, **36**, 960–972.
11. Chaudhary, S., *J. Geol. Soc. India*, 1985, **26**, 345–349.
12. Manjrekar, V. D., Bandopadhyay, D. N. and Ghosh, A., *Geophytology*, 1956, **16**, 145–152.
13. Chandra, S., *Indian J. Geol.*, 1989, **61**, 30–40.
14. Banerjee, M., Basu, M., Haldar, A. and Hait, A., *Indian Biol.*, 1991, **23**(2), 1–7.

KAMAL JEET SINGH
SHAILA CHANDRA

Birbal Sahni Institute of Palaeobotany,
53, University Road,
Lucknow 226 007, India.

Role of phospholipase C in the cryptobiotic cysts of the fairy shrimp *Streptocephalus dichotomus*

As our earlier report[1] disclosed the occurrence and functional significance of the enzyme PC-PLC (phosphatidylcholine cholinephosphohydrolase) in the cryptobiotic cysts of the fairy shrimp *Streptocephalus dichotomus*, the involvement of biological catalysts (enzymes) in the hydration-dependent hatching of the cysts becomes quite evident.

Ever since the introduction of term 'cryptobiosis' in biology by Keilin[2], it has fascinated many investigators due to its unique features like maintenance of embryonic viability during thermal stress and the ability to withstand dehydration. *S. dichotomus*, which inhabits the freshwater ponds, undergoes cryptobiosis during dry summer period in order to overcome the adverse conditions. Though a detailed study on the egg morphology[3], biochemistry[4] and reproductive biology[5]

has been carried out, the role of enzymes involved in signal transduction has not yet been investigated in detail.

Although PI-PLC (phosphatidylinositol phosphodiesterase) has been studied extensively in mammals and micro-organisms[6], reports on the role of its substrate, namely phosphoinositides, are available only in the embryonic development of sea urchins[7]. As a growing body of evidences disclose the significance of inositol phospholipids[8] and their hydrolytic products like DAG[8] in the transduction of signals from external to internal cellular system, the role of PI-PLC becomes seldom negligible.

As cyst hatching involves hydration for a definitive period (24 h), presumably some signal is transduced through messenger molecules. This, and the earlier results which revealed the role of calcium[9], $pH_{(i)}$ (ref. 10) and PC-PLC[1] initiated studies on whether or not the PI-PLC has any role in cyst hatching. Cysts were collected from the animals as described previously[1], homogenized in Tris–acetate buffer (0.01 m, pH 7.5) and centrifuged at $4500 \times g$ for 20 min in a KUBOTA-1500 refrigerated centrifuge. The supernatant was decanted and the pellet was redissolved in the same buffer and centrifuged as above. Both the supernatants were pooled up and used as the sample for the assay.

Activity assay of PI-PLC was carried out by the alkaline phosphatase releasing method as described by Ikezawa and Taguchi[11]. In brief, male rat (Albino, Wistar strain) kidneys were excised and homogenized in ice-cold 0.25 M sucrose solution. To the homogenate 0.5 ml of the sample was added and incubated at 37°C for 30 min. After the incubation, 1 ml aliquot was withdrawn, cooled in an ice bath and then centrifuged at $1,09,000 \times g$ for 90 min in a Beckman L7-65 ultracentrifuge. The supernatant represented the cytosolic fraction and the sediment the particulate fraction. Both cytosolic and particulate fractions were assayed for the alkaline phosphatase activity according to Engstrom[12]. The absorbance was recorded at UV–visible 160 Å spectrophotometer.

Figure 1 depicts the activity assay results of PI-PLC at different hours of hydration of the cysts of *S. dichotomus*. The enzyme activity of the cytosolic fraction increased in the first 12 h of hydration (0–12 h), followed by a subsequent decrease in the next 12 h (12–24 h) of hydration, whereas in the particulate fraction the activity of the enzyme decreased initially (0–12 h) and increased in the next 12 h (12–24 h) of hydration. Although in both cytosolic and particulate fractions the activity of the enzyme varied considerably between the first 12 h (0–12 h) and the next 12 h (12–24 h), a unique feature observed is that the activity in cytosolic fraction is inversely proportional to particulate fraction. In other words, whenever there is an increase/decrease in the PI-PLC activity of the cytosolic fraction, it is inversely proportional to the particulate fraction.

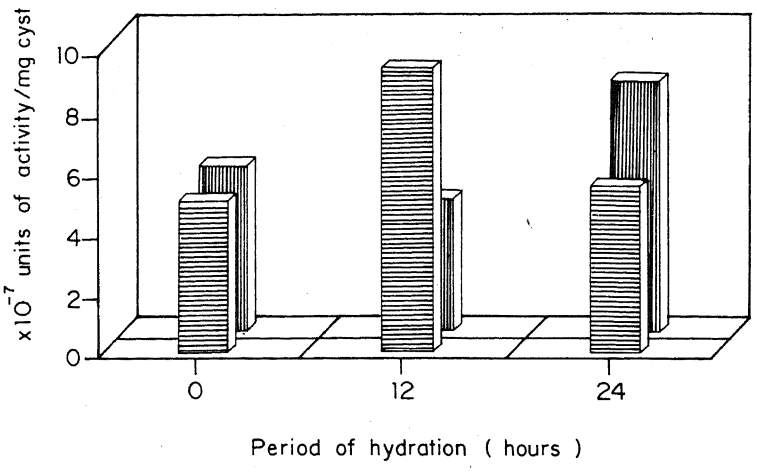


Figure 1. PI-PLC activity in the cryptobiotic cysts of the fairy shrimp Streptocephalus dichotomus. ▤, Cytosolic fraction; ▥, Particulate fraction.

Unlike PC-PLC, where there is a steady-state increase in the activity up to the period of hatching[1] (i.e. 24 h), PI-PLC has a fluctuating trend. As PI-PLC is a membrane-bound enzyme, it is presumable that the activity is high in the particulate fraction than in the cytosolic fraction at 0 h of hydration. The increase in activity in the cytosolic fraction at 12 h of hydration can presumably be accounted for its transmembrane mobility as it has to hydrolyse the cytosolic IP. Similarly, the decrease in activity in the cytosol at 12–24 h of hydration implies the culmination of PI-PLC-catalysed reaction.

From the present report it is evident that a change (increase/decrease) in the activity of PI-PLC does occur in both cytosolic and particulate fractions pertaining to the period of hydration.

1. Shanmugasundaram, G. K. and Munuswamy, N., *Biochem. Mol. Biol. Int.*, 1995, **36**, 203–207.
2. Keilin, D., *Proc. Roy. Soc. Lond. Series B*, 1959, **150**, 149–191.
3. Munuswamy, N. and Subramonium, T., *Cytobios.*, 1983, **37**, 181–186.
4. Bernice, R., *Hydrobiologia*, 1972, **39(2)**, 155–164.
5. Munuswamy, N. and Subramonium, T., *Crustaceana*, 1985, **49(2)**, 113–118.
6. Rhee, S. G., Suh, P. G., Ryu, S. H. and Lee, S. Y., *Science*, 1989, **244**, 546–550.
7. Turner, P. R., Sheetz, M. P. and Jaffe, L. A., *Nature*, 1984, **310**, 414–415.
8. Berridge, M. J, *Ann. Rev. Biochem.*, 1987, **50**, 159–193.
9. Dumont, H. J, Casier, P., Munuswamy, N. and De Walsche, *Hydrobiologia*, 1992, **230**, 1–7.
10 Munuswamy, N., Abdul Nazar, A. K., and Dumont, H. J., *Curr. Sci.*, 1994, **62**, 751–752.
11. Ikezawa, H. and Taguchi, R., *Methods Enzymol.*, 1981, **71**, 731–741.
12. Engstrom, L., *Biochim. Biophys. Acta*, 1964, **92**, 71–78.

ACKNOWLEDGEMENT. We gratefully acknowledge the CAS in Botany, University of Madras for providing the ultracentrifugal facility.

G. K. SHANMUGASUNDARAM
N. MUNUSWAMY

Department of Zoology,
University of Madras,
Guindy Campus,
Madras 600 025, India.

Reflecting diffusions*

S. Ramasubramanian

Probabilistic models of diffusion processes with 'reflecting boundary conditions' are discussed; the domains considered can be nonsmooth.

THE phenomenon of diffusion in free space (in a dilute solution) can be modelled in a mathematically rigorous fashion by probabilistic methods; in fact, this has been one of the major achievements of probability theory. This has led to a very fruitful interaction between the theory of partial differential equations and the theory of Markov processes; the interface is now known among probabilists as the theory of diffusion processes[1–5].

The aim of this exposition is to indicate how probabilistic methods are applied to model diffusions with 'reflecting boundary conditions'. To keep the length of the article within reasonable limits, we do not consider asymptotic properties, connections with boundary value problems, etc. And for the same reason it is inevitable that we indulge in a bit of probabilistic jargon without explanation; see refs 1, 3, 6–8 for more information concerning unexplained technical terms like submartingales, Markov property, stochastic integrals, . . .

Diffusions in $\mathbb{R}^d$

We now briefly review three ways of modelling diffusion processes in $\mathbb{R}^d$. This will help the reader appreciate the approaches taken later to study reflecting diffusions. Our notation will also get fixed in the process.

Let the operator A be defined by

$$Af(x) = \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(x)\frac{\partial^2 f(x)}{\partial x_i \partial x_j} + \sum_{i=1}^{d} b_i(x)\frac{\partial f(x)}{\partial x_i}, \qquad (1)$$

where the coefficients a_{ij}, b_i are sufficiently smooth real-valued functions on $\mathbb{R}^d$; A is assumed to be uniformly elliptic; $b(\cdot)$ is the *infinitesimal drift* and $a(\cdot)$ the *infinitesimal dispersion.*

Semigroup approach

Let $p(s, x; t, z)$, $0 \le s < t$, $x, z \in \mathbb{R}^d$ denote the (minimal) fundamental solution for the *Fokker–Planck equation*

$$\frac{\partial u}{\partial t}(t, z) = A_z^* u(t, z), \quad t > 0, \quad z \in \mathbb{R}^d, \qquad (2)$$

where A^* is the formal adjoint of A, and the subscript z denotes differentiation in the z variables; that is, if the initial value $u(0, \cdot) = f(\cdot)$ is specified then u can be given by

$$u(t, z) \equiv (T_t^* f)(z) = \int_{\mathbb{R}^d} f(x)p(0, x; t, z)\,dx. \qquad (3)$$

Since the concentration of a diffusing species satisfies the Fokker–Planck equation, knowledge of p determines the system completely once the initial concentration is known.

It can be shown that $p(s, x; t, \cdot)$ is a probability density function, satisfies the so-called Chapman–Kolmogorov equation, etc.; in other words, p is a transition probability density function (in the sense that $p(s, x; t, z)\,dz$ gives the probability that a diffusing particle starting from x at time s reaches a sufficiently small neighbourhood of z at time t). Hence, by the theory of Markov processes, there is a continuous-time Markov process in $\mathbb{R}^d$ with continuous sample paths for which p is the transition density. This process can be called *diffusion with generator A*. As p satisfies the Chapman–Kolmogorov equation, $\{T_t^* : t \ge 0\}$ given by eq. (3) forms a semigroup of operators. This is the classical approach to diffusions, pioneered by Kolomogorov and Feller, and depends on PDE theory for the guaranteed existence of the fundamental solution.

If $a(\cdot) \equiv (d \times d)$ identity matrix, $b(\cdot) \equiv 0$, then $A = (1/2)\Delta = A^*$, eq. (2) is the heat equation, p is the heat kernel and the corresponding diffusion is, of course, the standard Brownian motion (this explains the fondness for the factor 1/2 in eq. (1) among probabilists!). This has been the prototype of diffusion and corresponds to the motion of a diffusing particle purely due to fluctuations caused by bombardment by solvent particles in the absence of external forces, friction, etc.[8].

SDE approach

Levy had suggested that the motion of a diffusing particle can be represented, in differential form, by

$$dY(t) = \sigma(Y(t))dB(t) + b(Y(t))\,dt, \qquad (4)$$

S. Ramasubramanian is in the Indian Statistical Institute, 8th Mile, Mysore Road, Bangalore 560 059, India.

*This is a slightly enlarged verion of a lecture given at the Department of Mathematics, Indian Institute of Science, Bangalore on 10 February 1995.

where $\sigma(x)$ is a matrix such that $\sigma(x)\sigma^*(x) = a(x)$, $\{B(t): t \geq 0\}$ is a standard d-dimensional Brownian motion. The idea is: in a short interval of time the diffusing particle behaves locally as a Brownian particle with appropriate 'standard deviation' σ subject to an 'infinitesimal drift' b. However, it is well known that the Brownian trajectories are nowhere differentiable. It has been a seminal achievement of K. Ito to give meaning for integrals of the form $\int f(s, \omega)\, dB(s, \omega)$ for a large class of processes f; such integrals are called Ito integrals. Then the formal differential expression (4) can be interpreted as the stochastic integral equation

$$Y(t) = Y(0) + \int_0^t \sigma(Y(s))\, dB(s) + \int_0^t b(Y(s))\, ds. \qquad (5)$$

When σ, b are Lipschitz-continuous and satisfy certain growth conditions, it was shown by Ito (essentially by Picard's iteration) that eq. (5) has a unique solution $Y(\cdot)$; also the solution Y is a continuous Markov process. If Y has a sufficiently smooth transition density p, then it can be proved that p is the minimal fundamental solution for eq. (2). Thus, the solution of the *stochastic differential equation* (4) can be called[3, 9] diffusion with generator A.

When $d = 1$, $a \equiv 1$, $b(x) = -\beta x$ with $\beta > 0$, we get the Ornstein–Uhlenbeck process.

Martingale problem

An important feature of Ito integrals is Ito's formula for transformation of variables: for any $f \in C_b^2(\mathbb{R}^d)$,

$$f(Y(t)) - f(Y(0)) - \int_0^t Af(Y(s))\, ds = \int_0^t \langle \sigma(Y(s))\nabla f(Y(s)), dB(s) \rangle. \qquad (6)$$

If σ is bounded, the Ito integral on the right-hand side of eq. (6) is a martingale.

Since Y is a continuous process it induces a distribution (probability measure) on the path space $\Omega = C([0, \infty) : \mathbb{R}^d)$ = the space of $\mathbb{R}^d$-valued continuous functions on $[0, \infty)$. Denote this distribution by P_x, where x represents the starting point (that is, $Y(0) \equiv x$). Let $X(t, \omega) = \omega(t)$, $t \geq 0$, $\omega \in \Omega$. Then eq. (6) can be expressed as

$$f(X(t)) - f(x) - \int_0^t Af(X(s))\, ds = P_x - \text{martingale} \qquad (7)$$

for any $f \in C_b^2(\mathbb{R}^d)$.

The above can be used to characterize a diffusion process. The point of view taken is that the family $\{P_x : x \in \mathbb{R}^d\}$ of probability measures on Ω contains all relevant information about the diffusion process. In their fundamental work Stroock and Varadhan[11] have shown that, if a, b are bounded, continuous, then for each x there is a unique P_x such that eq. (7) holds for any $f \in C_b^2(\mathbb{R}^d)$; also under $\{P_x\}$ the coordinate projections $\{X(t) : t \geq 0\}$ form a continuous Markov process. Moreover, if the coefficients a, b are sufficiently smooth then $\{P_x\}$ agrees with the distribution of diffusion as defined earlier. Thus, the family $\{P_x : x \in \mathbb{R}^d\}$ of probability measures solving the martingale problem given by eq. (7) can be called diffusion process corresponding to A.

One may consult refs 3, 9, 11 concerning the advantages of the various approaches.

Two models

Insulated heat conduction

The problem of heat conduction in a semi-infinite bar (of uniform cross-section) with the end kept insulated is described by

$$\frac{\partial u}{\partial t}(t, z) = \frac{1}{2}\frac{\partial^2 u}{\partial z^2}(t, z), \quad t > 0, \ z > 0 \qquad (8)$$

subject to the boundary condition

$$\frac{\partial u}{\partial z}(t, 0) = 0, \quad t > 0. \qquad (9)$$

This is the one-dimensional heat equation with the Neumann condition at $z = 0$. Equations (8) and (9) describe also the concentration of a species undergoing Brownian movement (confined to $[0, \infty)$) in a dilute solution when no mass is allowed to 'escape through' 0 or to 'linger at' 0. It is well known that the fundamental solution for the above is given by

$$q(s, x; t, z) = \frac{1}{\sqrt{2\pi(t-s)}} \times \left[\exp\left\{-\frac{(z-x)^2}{2(t-s)}\right\} + \exp\left\{-\frac{(-z-x)^2}{2(t-s)}\right\}\right] \qquad (10)$$

for $0 \leq s < t$, $x \geq 0$, $z \geq 0$.

By arguments analogous to those presented earlier, it can be shown that there is a Markov process (living in $[0, \infty)$) with q as its transition probability density. This process may be called *reflected Brownian motion* (RBM). If $\{B(t)\}$ is a one-dimensional Brownian motion (in $\mathbb{R}$), then it is easily seen that $\{|B(t)| : t \geq 0\}$ is a Markov process with q as its transition density; that is, $\{|B(t)| : t \geq 0\}$ is RBM.

An inventory model

Consider the following model in queueing theory:

$$Z(t) = B(0) + I(t) - [O(t) - L(t)] = B(t) + L(t), \qquad (11)$$

$t \geq 0$, representing the *inventory process* Z in a simple two-stage flow system consisting of input (or production), output (or demand) and an intermediate buffer storage of infinite capacity. Here $I(t)$ = cumulative input up to t; $O(t)$ = cumulative potential output (or demand) up to t, $I(0) = O(0) = 0$ and $B(0) \geq 0$ is the initial inventory level. It is assumed that the demand that cannot be met from the stock on hand is lost with no adverse effect on future demand; thus, $L(t)$ can be interpreted as the demand that could not be met up to t, and $O(t)$–$L(t)$ is the actual output over $[0, t]$. Clearly, $Z(t) \geq 0$. All are assumed to be continuous stochastic processes; L is called *lost demand process* and B the *netput process*[12].

Under certain conditions of *heavy traffic*, $\{B(t)\}$ can be modelled by one-dimensional Brownian motion: viz. (a) in any short time interval $[t_1, t_2]$, the input $I(t_2)$–$I(t_1)$ and the output $O(t_2)$–$O(t_1)$ are quite large but their difference (netput) $B(t_2)$–$B(t_1)$ is not large; (b) netput increments over nonoverlapping intervals are approximately independent.

With these assumptions the set-up (11) becomes

$$Z(t) = B(t) + L(t), \quad t \geq 0, \tag{12}$$

satisfying (i) $Z(t) \geq 0$, Z is continuous in t, (ii) $Z(0) = B(0) \geq 0$, (iii) $L(0) = 0$, L is nondecreasing and continuous in t, (iv) L increases only when $Z = 0$, (v) $\{B(t)\}$ is a Brownian motion.

A look at the above indicates that a 'deterministic version' of eq. (12) with the attendant constraints is meaningful. To determine z and l such that

$$z(t) = \alpha(t) + l(t), \quad t \geq 0, \tag{13}$$

subject to (i)′–(iv)′, where (i)′–(iv)′ are the same as (i)–(iv) with l, z, α replacing, respectively, L, Z, B, and (v)′ $\alpha(\cdot)$ is a continuous function. (Here α is the known function). For a continuous function α on $[0, \infty)$ with $\alpha(0) \geq 0$ the deterministic problem is solved uniquely by taking

$$l(t) = -\inf\{\alpha(s) \wedge 0 : 0 \leq s \leq t\} \tag{14}$$

and, of course, $z(t) = \alpha(t) + l(t)$. A hard stare at Figure 1 will convince the reader!

Consequently, the stochastic problem can also be uniquely solved path by path. This problem is called *Skorohod problem* and eq. (12) is called *Skorohod equation.*

The connection and the consequences

What do the above-described models have in common? According to Levy[3], practically everything!

$\{|B(t)| : t \geq 0\}$ and $\{Z(t) : t \geq 0\}$, being continuous processes, induce corresponding distributions in the path space $C([0, \infty) : \mathbb{R})$. Levy has shown that these distributions are identical if the initial values agree. This surprising result is proved[3] by showing that Z is a Markov process with transition density q given by eq. (10). It may be noted that Levy arrived at this result purely by studying the sample path properties of Brownian motion, without any queueing theoretic considerations.

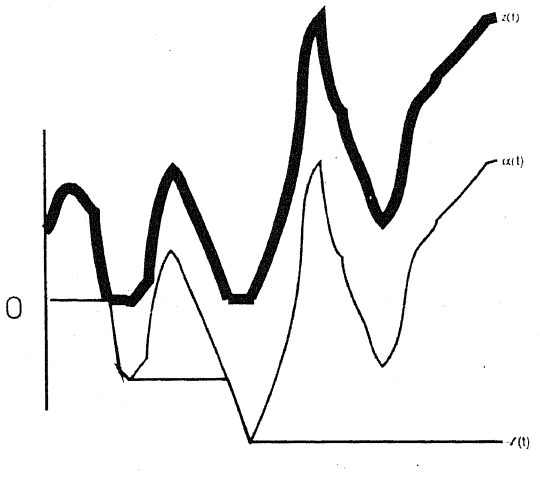


Figure 1.

Since L increases only when Z is on the boundary, L is called the *boundary local time* of Z. Local time process is one of the most esoteric objects associated with Brownian motion; in a sense high irregularity of Brownian paths makes the local time a reasonably well-behaved entity[3].

To indicate how the local time and transition density are tied up, we consider the RBM in the half-space $D = \{x \in \mathbb{R}^d : x_1 \geq 0\}$ with reflection in the inward normal direction; note that the inward normal derivative is just $\partial/\partial x_1$ at the boundary.

Let $\{(B_1(t), B_2(t), \ldots, B_d(t)) : t \geq 0\}$ be the d-dimensional Brownian motion in $\mathbb{R}^d$ (without any boundary condition). From the preceding discussion RBM in D with normal reflection at the boundary can be represented as

$$Z_1(t) = B_1(t) + L(t), \quad Z_i(t) = B_i(t), \quad i = 2, \ldots, d, \tag{15}$$

where $L(t) = -\inf\{B_1(s) \wedge 0 : 0 \leq s \leq t\}$. Also the transition probability density is given by

$$p(s, x; t, z) = q(s, x_1; t, z_1)\ r(s, \bar{x}; t, \bar{z}), \tag{16}$$

where $0 \leq s < t$, $x = (x_1, x_2, \ldots, x_d) = (x_1, \bar{x}) \in D$, $z = (z_1, z_2, \ldots, z_d) = (z_1, \bar{z}) \in D$, q is given by eq. (10) and r is the $(d-1)$-dimensional heat kernel. Using the Ito calculus and integration by parts it can be shown that

$$E_x\left[\int_0^t g(Z(r))\, dL(r)\right] = \frac{1}{2}\int_0^t \int_{\partial D} g(z) p(0, x; \alpha, z)\, d\sigma(z)\, d\alpha \tag{17}$$

for any $x \in D$ and bounded measurable function g on ∂D; here $d\sigma(\cdot)$ denotes the surface area measure on the

boundary (in this case, $(d-1)$-dimensional Lebesgue measure) and E_x denotes taking expectation under the condition $Z(0) \equiv x$.

The case of smooth domains

The importance of the Skorohod equation is that it suggests how reflecting diffusion can be represented as a solution of a stochastic differential equation. Also one may consider reflecting directions other than the normal direction, and possibly depending on $x \in \partial D$.

Let D be a domain with smooth boundary and $\gamma(\cdot)$ a smoothly varying vector field on ∂D; assume that $\langle \gamma(x), n(x) \rangle \geq \beta > 0$, $x \in \partial D$, where $n(x)$ is the inward normal. *Reflecting diffusion* in D with generator A and reflecting field γ can be defined as the process $\{Z(t)\}$ satisfying

$$Z(t) = Z(0) + \int_0^t \sigma(Z(s))\, dB(s) + \int_0^t b(Z(s))\, ds + \int_0^t \gamma(Z(s))\, d\xi(s) \qquad (18)$$

such that $\xi(0) = 0$, ξ is a continuous, nondecreasing process, ξ increases only when Z is on the boundary, and $Z(t) \in \overline{D}$ for all t. This formulation is due to Watanabe[9]. Note, however, the difference between eqs (4) and (18); in eq. (4) there is only one unknown process where (18) involves two unknown processes Z and ξ. It can be shown that eq. (18) has a unique solution which is Markov. In addition, if Z has a sufficiently regular transition probability density p then

$$\frac{\partial p}{\partial s}(s, x; t, z) + A_x p(s, x; t, z) = \delta(t-s)\delta(z-x), \quad s < t, \ x \in D \qquad (19)$$

$$\langle \gamma(x), \nabla_x p(s, x; t, z) \rangle = 0, \quad x \in \partial D \qquad (20)$$

for fixed t, z. (Observe that eq. (19) is the adjoint of the Fokker–Planck equation, known as the *backward Kolmogorov equation*.) In such a case the analogue of eq. (17) also holds. It may be noted that the notion 'reflecting direction' is justified mainly in view of eq. (20). As the sample paths of a diffusion process are nowhere differentiable, it is meaningless to say that a trajectory gets reflected in a particular direction; however, the boundary is 'reflecting' in the sense of Markov processes, viz. the sample path does not linger on the boundary and it is instantaneously returned to the interior.

Submartingale problem

As in the case of diffusion in $\mathbb{R}^d$, reflecting diffusions in smooth domains can also be characterized as solutions of 'submartingale problem', a development due to Stroock and Varadhan[10] once again. That is, for each $x \in \overline{D}$ we seek a probability measure P_x on the path space such that

(i) $P_x(X(t) \in \overline{D} \ \ \forall t \geq 0, \ X(0) = x) = 1$,

(ii) for each $f \in C_b^2(\mathbb{R}^d)$ with $\langle \gamma, \nabla f \rangle \geq 0$, on ∂D,

$$f(X(t)) - f(x) - \int_0^t Af(X(s))\, ds = P_x - \text{submartingale}, \qquad (21)$$

where X is as defined earlier. If a, b are bounded, continuous and γ bounded, Lipschitz-continuous, then the submartingale problem has a unique family $\{P_x : x \in \overline{D}\}$ of solutions; in such a case, under $\{P_x\}$ the coordinate projections $\{X(t)\}$ form a continuous Markov process and

$$E_x\left[\int_0^\infty I_{\partial D}(X(s))\, ds\right] = 0, \quad x \in \overline{D}, \qquad (22)$$

where E_x denotes the expectation with respect to P_x; eq. (22) means that the time spent on ∂D is zero. When a, b, γ are sufficiently smooth, $\{P_x\}$ agrees with the distribution of the reflecting diffusion. Thus, we have another way of looking at reflecting diffusions: the family $\{P_x : x \in \overline{D}\}$ of probability measures solving the above submartingale problem can be called[10] the reflecting diffusion process corresponding to A, γ.

Another inventory model

Consider the three-stage flow system (also called the *tandem buffer system*) outlined in Figure 2. Each buffer satisfies the earlier assumptions. Here $I_1(t)$ = cumulative input into buffer 1 up to t, $I_2(t)$ = potential transfer from buffer 1 to buffer 2 up to t = potential output from buffer 1 (potential input into buffer 2) up to t, $I_3(t)$ = potential output from buffer 2 up to t. The netput process (B_1, B_2) is defined by

$$B_1(t) = B_1(0) + I_1(t) - I_2(t),$$
$$B_2(t) = B_2(0) + I_2(t) - I_3(t). \qquad (23)$$

Let $Z_i(t)$ = content of buffer i at time t, $i = 1, 2$. We may write $Z_1(t) = B_1(t) + L_1(t)$, where $L_1(t)$ = amount of potential output lost up to t because of buffer 1 being empty. Consequently, $I_2(t) - L_1(t)$ = cumulative input into buffer 2 up to t. Therefore, applying a similar argument to buffer 2 we get $Z_2(t) = B_2(t) - L_1(t) + L_2(t)$, where $L_2(t)$ = amount of potential output lost up to t because of buffer 2 being empty. Clearly, $Z_i \geq 0$.

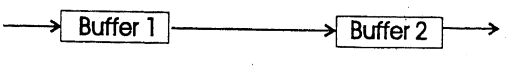


Figure 2.

In the case of heavy traffic, when the netput process can be assumed to be a two-dimensional Brownian motion $\{(B_1(t), B_2(t))\}$, the above suggests the problem: to find continuous processes $Z = (Z_1, Z_2)$, $L = (L_1, L_2)$ such that for $t \geq 0$,

$$Z_1(t) = B_1(t) + L_1(t) \geq 0,$$
$$Z_2(t) = B_2(t) - L_1(t) + L_2(t) \geq 0, \tag{24}$$

with $Z_i(0) = B_i(0) \geq 0$, L_i continuous, nondecreasing; L_i increases only when $Z_i = 0$, $i = 1, 2$. Note that Z takes values in the positive quadrant $\{(z_1, z_2): z_1 \geq 0, z_2 \geq 0\}$; the 'direction of reflection' on the boundary $\{z_2 = 0\}$ is the normal direction, whereas on $\{z_1 = 0\}$ it is the 'diagonal' direction $(-1, 1)$.

The above model is due to Harrison[12] and has given a lot of impetus to the study of reflecting diffusions in nonsmooth domains, especially to RBM in quadrant, wedge, orthant, etc. with oblique reflection at the boundary. Mathematical interest is due to the presence of corners and the reflecting direction field having 'discontinuity', and the consequent nonavailability of suitable results from the theory of partial differential equations. Here 'discontinuity' is to be understood as the difference between the reflecting direction and the normal direction being discontinuous.

Normal reflection in nonsmooth domains

What is meant by normal direction at a boundary point when the boundary is nonsmooth? Following Tanaka[13] and Lions and Sznitman[14], Saisho[15] has taken the approach given below.

For a domain D, $x \in \partial D$, $r > 0$, put $\mathcal{N}(x, r) = \{n \in \mathbb{R}^d : |n| = 1, B(x - rn, r) \cap D = \emptyset\}$, where $B(y, r)$ denotes the ball with centre y and radius r; the set $\mathcal{N}(x) = \cup_{r>0} \mathcal{N}(x, r)$. $\mathcal{N}(x)$ is defined to be the *set of all inward normal unit vectors* at x.

If ∂D is smooth at x, than $\mathcal{N}(x) = \{n(x)\}$. When D is the positive quadrant, $\mathcal{N}((0, 0)) = \{(x_1, x_2) : x_1 \geq 0, x_2 \geq 0, \quad x_1^2 + x_2^2 = 1\}$. Let $D = \{(-1, 0) \times (-1, 1)\} \cup \{(0, 1) \times (-1, 1)\}$; that is, D is a rectangle with a 'fibre' removed; note that for any x on the fibre $\mathcal{N}(x) = \emptyset$.

If D satisfies a uniform exterior sphere condition then $\mathcal{N}(x) \neq \emptyset$ for $x \in \partial D$; for such domains the deterministic Skorohod problem can be formulated as follows: Let $\alpha \in C([0, \infty) : \mathbb{R}^d)$ with $\alpha(0) \in \overline{D}$. Find a pair of continuous functions z, k such that

$$z(t) = \alpha(t) + k(t) \in \overline{D}, \quad k(t) = \int_0^t n(s)\, d|k|(s), \tag{25}$$

where $|k|(t)$ is continuous, nondecreasing, and increases only when z is on the boundary ∂D, and $n(s) \in \mathcal{N}(z(s))$ if $z(s) \in \partial D$. Note that k is $\mathbb{R}^d$-valued and is of bounded variation in every bounded interval.

Under a further technical assumption which may be called a uniform interior cone condition, it has been shown by Saisho[14], based on the work of Lions and Sznitman[16], that a unique pair of solutions exists for any α.

Moreover, if σ, b are bounded, Lipschitz-continuous, $Z(0) \in \overline{D}$, and $\{B(t)\}$ is a d-dimensional Brownian motion, then there exist unique continuous processes $\{Z(t)\}$, $\{K(t)\}$ such that Z is $\overline{D}$-valued, K is $\mathbb{R}^d$-valued with bounded variations,

$$Z(t) = Z(0) + \int_0^t \sigma(Z(s))\, dB(s) + \int_0^t b(Z(s))\, ds + K(t), \tag{26}$$

$$K(t) = \int_0^t n(s)\, d|K|(s), \tag{27}$$

where $|K|(t)$ is a continuous, nondecreasing process, and increases only when Z is on the boundary, and $n(s) \in \mathcal{N}(Z(s))$ if $Z(s) \in \partial D$. The process Z may be called reflecting diffusion in $\overline{D}$ with generator A and normal reflection at the boundary[13–15].

If Z is a quadrant/rectangle/orthant, $\sigma \equiv$ identity matrix, $b \equiv 0$, then RBM with normal reflection can be easily got from the d-dimensional Brownian motion by the method of images; in such a case Z constructed above agrees with RBM in law.

RBM with oblique reflection

Heavy-traffic approximation in queueing theory has led in recent years to an intensive study of RBM in domains like quadrant/wedge/orthant with oblique reflection. Not much is known about other diffusion processes; and even about RBM the last word is not said. We will confine ourselves to a brief discussion of RBM in a wedge; because of the 'discontinuity' at the corner, there are a few surprises around the corner!

Let D denote the wedge with angle ξ; $\partial_1 D$, $\partial_2 D$ are the two sides of the wedge. We assume that the direction of reflection is constant on each side. Let v_i denote the reflecting direction on $\partial_i D$, θ_i denote the angle v_i makes with the normal, $i = 1, 2$ (θ_i is taken to be positive if v_i points towards the corner) (Figure 3).

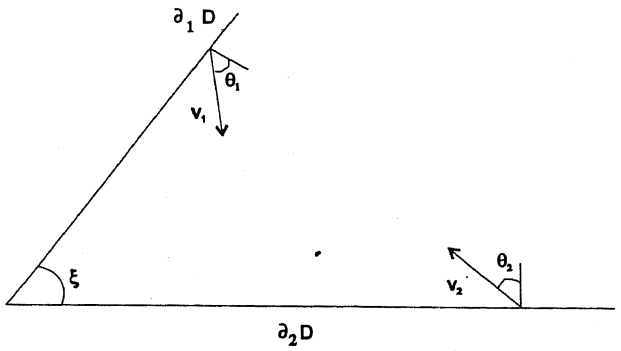


Figure 3.

In their fundamental paper Varadhan and Williams[16] defined RBM in a wedge as the solution to the following submartingale problem. Let $\mathfrak{I}$ be the collection of all smooth functions f on $\mathbb{R}^2$ such that f is constant near $(0, 0)$ and $\langle v_i, \nabla f\rangle \geq 0$ on $\partial_i D$, $i = 1, 2$. Let Ω, X be as before. A family $\{P_x : x \in \overline{D}\}$ of probability measures on Ω is a solution to the submartingale problem corresponding to $((1/2)\Delta, v_1, v_2)$ if the following hold for each $x \in \overline{D}$:

(i) $P_x(X(t) \in \overline{D} \ \ \forall t \geq 0, \ X(0) = x) = 1$;

(ii) for any $f \in \mathfrak{I}$

$$f(X(t)) - f(x) - \int_0^t \frac{1}{2}\Delta f(X(s))\, ds = P_x - \text{submartingale}; \tag{28}$$

(iii) $$E_x \int_0^\infty I_{\{0\}}(X(s))\, ds = 0, \tag{29}$$

where $\Delta = \partial^2/\partial x_1^2 + \partial^2/\partial x_2^2$. Set $(\theta_1 + \theta_2)/\xi$. The parameter α determines crucially the behaviour of the submartingale problem[16], viz.:

(a) $\alpha \leq 0$. There is a unique solution to the problem, and the corner is never visited.
(b) $0 < \alpha < 2$. There is a unique solution. The process hits the corner with probability one, instantaneously leaves the corner.
(c) $\alpha \geq 2$. There is *no* solution satisfying (i)–(iii) above. However, for each $x \in \overline{D}$ there is a unique P_x satisfying (i) and (ii); this P_x is supported on those paths which reach the corner and terminate there; that is, the process is *absorbed* at the corner.

Note that if $\theta_i = 0$, $i = 1, 2$, we have the case of normal reflection, and the solution of the submartingale problem is the distribution of the process considered in the previous section with, of course, $\sigma \equiv$ identity matrix, $b \equiv 0$.

The condition $\alpha > 0$ roughly means that there is a net push towards the corner. In such a case the corner is hit, and if the push is beyond a critical level (viz. $\alpha \geq 2$), the corner becomes an absorbing barrier! This is a surprising development as diffusions in $\mathbb{R}^d$ and in smooth domains (with $d \geq 2$) do not hit any specified point.

Proof of the result of Varadhan and Williams[16] centres around the function

$$G(r, \theta) = \begin{cases} r^\alpha \cos(\alpha\theta - \theta_2), & \alpha \neq 0, \\ \log r + \theta \tan\theta_2, & \alpha = 0. \end{cases} \tag{30}$$

which is the function (in polar coordinates) satisfying $\Delta G = 0$ in D, $\langle v_i, \nabla G\rangle = 0$ on $\partial_i D \backslash \{(0, 0)\}$, $i = 1, 2$.

Another surprising result, due to Williams[17], is that RBM in a wedge (that is, the solution of a submartingale problem) cannot be realized as a semi-martingale for $1 < \alpha < 2$; in other words, representation in terms of a stochastic differential equation (s.d.e) is not possible! This again is in contrast with diffusions in $\mathbb{R}^d$ or in smooth domains.

It may be mentioned that when D = quadrant, $\theta_i \geq 0$, $i = 1, 2$, $\theta_1\theta_2 < 1$, Harrison and Reiman[12, 17] have constructed RBM as the solution of an s.d.e by considering a modified Skorohod problem. Of late there has been considerable interest concerning the Skorohod problem with oblique reflection in nonsmooth domains.

For more information about RBM with oblique reflection the interested reader may refer to the forthcoming survey article by Williams[17].

1. Bhattacharya, R. N. and Waymire, E. C., *Stochastic Processes with Applications*, John Wiley, New York, 1992.
2. Freidlin, M., *Functional Integration and Partial Differential Equations*, Princeton Univ. Press, New York, 1985.
3. Karatzas, I. and Shreve, S. E., *Brownian Motion and Stochastic Calculus*, Springer, New York, 1991.
4. Taira, K., *Diffusion Processes and Partial Differential Equations*, Academic Press, New York, 1988.
5. Varadhan, S. R. S., *Diffusion Problems and Partial Differential Equations*, Narosa Publishing House, New Delhi, 1980.
6. Feller, W., *An Introduction to Probability Theory and its Applications*, Wiley-Eastern, New Delhi, 1971, vol. 2.
7. Karlin, S. and Taylor, H. M., *A First Course in Stochastic Processes*, and *A Second Course in Stochastic Processes*, Academic Press, New York, 1981.
8. Van Kampen, N. G., *Stochastic Processes in Physics and Chemistry*, North-Holland, Amsterdam, 1992.
9. Ikeda, N. and Watanabe, S., *Stochastic Differential Equations and Diffusion Processes*, North-Holland/Kodansha, Amsterdam/Tokyo, 1981.
10. Stroock, D. W. and Varadhan, S. R. S., *Commun. Pure Appl. Math.*, 1971, **24**, 147–225.
11. Stroock, D. W. and Varadhan, S. R. S., *Multidimensional Diffusion Processes*, Springer, New York, 1979.
12. Harrison, J. M., *Brownian Motion and Stochastic Flow Systems*, John Wiley, 1985.
13. Tanaka, H., *Hiroshima Math. J.*, 1979, **9**, 163–177.
14. Lions, P. L. and Sznitman, A., *Commun. Pure Appl. Math.*, 1984, **37**, 611.
15. Saisho, Y., *Probab. Theor. Rel. Fields*, 1987, **74**, 455–477.
16. Varadhan, S. R. S. and Williams, R. J., *Commun. Pure Appl. Math.*, 1985, **38**, 405–443.
17. Williams, R. J., Semimartingale reflecting Brownian motions in the orthant, *Preprint*, 1995.

Structural, clinical and transmission aspects of dioxins – Potential environmental pollutants

C. Ganesh Kumar and K. Jayaraj Rao

Dioxins are toxic and potent carcinogenic compounds formed as contaminants in milk and foods. Their occurrence, structure, prevalence in different food products, transmission to man and detection methods are discussed briefly. The dioxins and related compounds are found as trace contaminants in the synthesis of several commercial products, importantly chlorophenols. Dioxins have also been reported to be present in cigarettes, smoke and ash. The air-borne dioxins are transmitted to man either by direct inhalation or via forage, livestock, milk and meat. Mothers who have dioxins in the body could transfer them to the infants through breast milk, which may lead to late haemorrhagic disease in the newborn infants. It is important that there should be a check on the emission of 2,3,7,8-tetrachlorodibenzo-p-dioxins by industries, and pollution laws should take cognizance of these toxic compounds. This aspect is more relevant to India in view of the rapid industrialization taking place today.

IN the recent past, the environmental issues have become a matter of increasing concern in most of the countries. With ever-increasing human activity and industrialization, the number of pollutants to be dealt with are also increasing. Of these pollutants, dioxins – a family of compounds like polychlorinated dibenzo-*p*-dioxins (PCDDs) and polychlorinated dibenzofurans (PCDFs) – are known to be potent carcinogenic compounds and pose health risk to man. Of several dioxins and their congeners, 12 are of utmost importance because of their high toxicity.

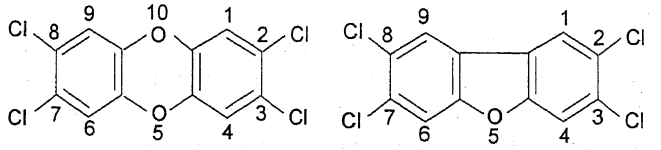
Dioxins are not synthesized commercially, but are formed as trace contaminants during the synthesis of several commercial products, mainly chlorophenols. For example, 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) is formed as a contaminant in the synthesis of 2,4,5-trichlorophenol, which is used in the manufacture of 2,4,5-trichlorophenoxy acetic acid (2,4,5-T), a broad spectrum herbicide and defoliant. Several other chlorophenols have wide usage as mold inhibitors, fungicides, insecticides and disinfectants. Of these, the most important one is pentachlorophenol, used as a fungicide for slime control during the manufacture of paper pulp, tanning of leather, etc.

Structure

The typical structure of a dioxin is shown in Figure 1. The dioxins have 75 possible chlorine-substituted bidenzo-*p*-dioxin isomers and 135 possible chlorine-substituted dibenzofuran congeners. The various isomers of dioxins formed due to the number and the position of the chlorine atoms are associated with the degree of toxicity. However, not all isomers are toxic.

Sources of dioxins

The major sources of these compounds are the impurities in chlorinated chemical wastes and the contaminants in ash and gas. Most of the sources are anthropogenic. Some PCDD and PCDF emissions arise from various combustion processes, including the incineration of municipal solid wastes (MSW) or hazardous wastes[1,2], power-producing stations using fossil fuels, and smaller systems such as private heating systems, fireplace and automobiles. Other industrial processes involved include the pulp and paper industry, steel mills, metal refining and the smelting industry. Releases are also reported to occur during hazardous waste accidents, illegal dumping, waste oil releases, and the emission of residues of treated wood, fly ash, and the production and application of pesticides[3]. There are genuine concerns that these emissions could also contaminate animal feeds and human foods such as milk, meat and vegetables[1,4]. The presence of these substances in the chlorine-bleached paper cartons used for milk packaging has also been reported[5,6]. The PCDDs and PCDFs could migrate from



2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) 2,3,7,8-tetrachloro-dibenzofuran (TCDF)

Figure 1. Structure of a dioxin and its congener.

C. Ganesh Kumar and K. Jayaraj Rao are in the Dairy Microbiology Division and Dairy Technology Division, respectively, National Dairy Research Institute, Karnal 132 001, India.

the bleached-paper cartons into the milk during storage. Among several factors which influenced the extent of this migration were the contamination level of the carton and the time of storage[7]. Reports reveal the presence of dioxins even in other food products like chicken and pork[8], fish and shell fish[9].

Permissible and prevalent levels of dioxins

In general, the dioxin content is represented by TCDD toxic equivalents used for assessing and comparing the dioxin and furan contents found in analyses. The proposed tolerable weekly intake of dioxins should not exceed 35 pg TCDD toxic equivalents per kg body weight, as recommended by Environmental Protection Agency (EPA) of the United States[10].

The dioxins have been detected frequently in milk, blood and adipose tissue and the concentrations varied in different tissues corresponding to the fat content. Cow milk is reported to contain 1.5 pg of dioxin/g fat[10]. Dioxins may find their way into the milk through cattle feed, however, some migration was also shown from milk cartons to the extent of 40 pg and 750 pg of TCCD and TCDF per kg fat, respectively. According to the estimates, the daily intake of TCDD toxic equivalents from three glasses of such contaminated milk exceeded the acceptable daily dose[11]. Comparatively, higher dioxin contents are reported in animal products. Over 60% of fat samples of chicken and pork in the study[8] contained more than 10 ng/g of pentachlorophenol (PCP). Of the poultry liver samples, 27% contained more than 50 ng of PCP per gram of liver. Moreover, PCP is known to contain a variety of chlorinated dibenzo-*p*-dioxins and chlorinated dibenzofurans[12].

There are alarming reports on the presence of chlorinated dioxins and dibenzofurans in human milk from several countries. Human milk may contain 10–50 times of these contaminants than cow milk or milk substitutes[13]. Analysis of human milk samples from South Africa, Pakistan, USA, Vietnam and England revealed dioxin toxicity levels as ranging from 8 to 23 ng/kg[14]. But, according to the analytical study conducted by the Federal Health Office (FHO), Berlin, the average daily intake of PCDD/PCDF by a baby was found to be 90 pg/kg body weight toxicity equivalents on a fat basis[15], which exceeds the permissible limits.

Transmission of dioxins to man

The potential source of human exposure to dioxins are phenoxy herbicides and airborne emissions resulting from combustion and waste disposal. The dioxins make an entry into human body through inhalation of contaminated air, but the main sources responsible for entry into human body are the milk and other food products. The dioxins are deposited on the forage/feed and soil.

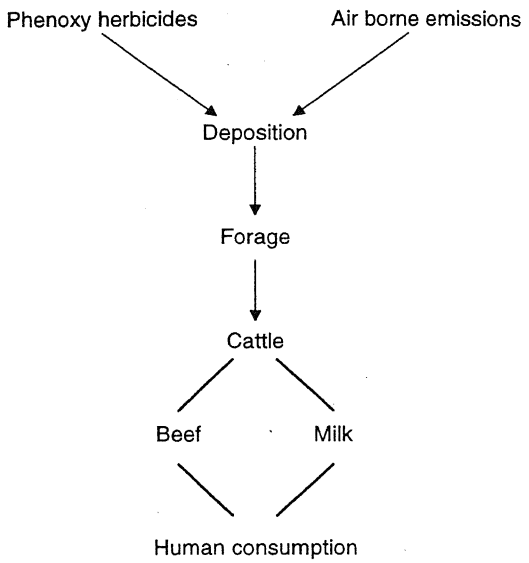


Figure 2. Flowchart for the transmission of dioxins to man.

They enter the body system of cattle when the latter ingest the contaminated feed[4, 17] or soil during grazing[16] (Figure 2). Of the total dioxins taken in by a cow, approximately 20% of the TCDD toxic equivalents are reported to be excreted in the milk[17].

Human population might be prone to dioxins where eating of smoked foods like sausages, fish and cheese is prevalent. It is important in this context to mention some of the smoked Indian dairy products like *Bandal* cheese and *Chhurpi*. *Chhurpi* is severely smoked and dried before selling in the market. This product is regularly consumed by tribal people and no studies as yet have been conducted to check possible dioxin transmission to these people. Another important source of dioxins may be the *khoa*, which is inadvertently smoked while being made over firewood in villages. This practice, however, is now discouraged and instead steam pans are used by the organized sector. Bluthgen and Heeschen[18] showed that 50% of average daily dioxin intake could be via milk products.

Studies relating to transmission of dioxins and their congeners have drawn much attention, especially with regard to transfer of these compounds from mother to child. The main factors influencing the transfer are the mother's dietary habits and personal habits such as cigarette smoking. The presence of dioxins in cigarette smoke has already been proved[19]. It could be said that in this modern society, more so in the western society, smoking even by pregnant women has become a fashion, which may lead to the passage of dioxins from the mother to child. Effects of passive smoking have also to be considered. Mothers who have dioxin content could transfer them to the infants through breast milk which may lead to late haemorrhagic disease in the newborn child[20]. In spite of these, encouragement of breast feed-

ing is recommended as its positive effects are considered more important than the risk of dioxins[21].

Clinical manifestations of dioxins

The dioxin 2,3,7,8-TCCD on long-term effect in small doses is carcinogenic, produces adverse reproductive effects and alters the body's immune response. Human beings exposed to large amounts of dioxins show many different symptoms like psychological disturbances, various sorts of physical disturbances, and neurological effects. The toxic syndromes produced by TCDD and related compounds[22] are:

1. Wasting syndrome.
2. Lymphoid ovulation – TCCD & congeners produce a loss of lymphoid tissue, especially in the thymus, but also in the spleen and lymph nodes.
3. Hepatotoxicity – TCCD and related compounds produce hepatotoxicity in all species even at low doses well below the lethal dose.
4. Gastric lesions.
5. Urinary tract hyperplasia. In the guinea pig, monkey and cow the transitional epithelial lining of the urinary tract (renal pelvis, ureter and urinary bladder) proliferates to 2–3 times its normal thickness.
6. Chloracne, a rare and potential debilitating disease, but important acneiform eruption is often associated with the ingestion of chlorinated phenolic agents such as dioxins.

Italian researchers have shown TCDD to produce chick oedema disease in fowls. Postmortem examination of the intoxicated fowls showed subcutaneous oedema, hyperpericardium, ascites, enlarged liver, oedema of myocardium, lungs and skeletal muscle, atropy of lymphoid tissue of the spleen and atropy of ovaries/testes[23,24].

Detection and quantification

Detection procedures vary from laboratory to laboratory and also with the type of sample to be analysed. The most commonly employed are chromatographic, spectrometric or immunological methods.

The dioxins and related compounds are soluble in organic solvents and hence are extracted from a product using acetone–hexane mixtures or hexane alone. These are then purified and chromatographed before determination by a combination of gas chromatography and mass spectrometry (GC–MS)[25]. Other techniques like capillary column high-resolution gas chromatography/high-resolution mass spectrometry (HRGC/HRMS), packed column gas low-resolution gas chromatography/high-resolution mass spectrometry (LRGC/HRMS) and radioimmunoassay have also been employed[26, 27]. The detection limits varied from 0.005 to 6 pg/g of the sample analysed.

Conclusions

Reports till date on dioxins and related compounds establish that these compounds could be a potential health risk in the long run if adequate precautionary measures are not taken. This is significant at a time when all the countries in the world are planning for a concerted action to check environmental pollution. Incidence of dioxins is high in industrialized countries and extensive work is being carried out in countries like USA, Germany, Sweden, Japan, etc. In the Asian context, India is the highest producer of livestock, and thus the dependence of humans on livestock and its products is high. These livestock are dependent on the soil and plant systems, which are being highly influenced by the 'modern agricultural systems'.

India is entering a fast growth phase of industrialization. The emissions from the chemical industries will result in the emission of PCBs, PCDDs and PCDFs in the surrounding environment, which will have an impact on the soil–plant systems of that area. It is desirable to assess the TCDD values in soil and crops around the sources of emission. Several studies have been attempted to determine the formation mechanism of PCDD/PCDF, but there still exist many areas where the information is incomplete.

Sufficient data are lacking in India to compare the levels of dioxin emission and consumption with the standards already established in developed countries. The standards and check levels of the industries around the farms should be introduced and implemented to the maximum level.

1. Travis, C. C. and Hattemer, F. H. A., *Risk Anal.*, 1989, **9**, 91–97.
2. Williams, P. T., *J. Institut. Water Environ. Manage.*, 1990, **4**, 26–34.
3. Hutzinger, O. and Fiedler, H., *Chemosphere*, 1989, **18**, 23–32.
4. Fries, G. F. and Paustenbach, D. J., *J. Toxicol. Environ. Health*, 1990, **29**, 1–43.
5. Anonymous, *Dairy Ind. Int.*, 1989, **54**, 42–43.
6. Paterson, K., *Milk Ind. London*, 1990, **92**, 26–27.
7. Rappe, C., Lindstrom, G., Glas, B., Lundstrom, K. and Borgstrom, S., *Chemosphere*, 1990, **20**, 1649–1656.
8. Ryan, J. J., Lizotte, R., Sakuma, T. and Mori, B., *J. Agric. Food Chem.*, 1985, **33**, 1021–1026.
9. Slorach, S. and Bergqvist, P. A., *Var Foda*, 1988, **40**, 437–444.
10. Rappe, C., *Scand. Dairy Inform.*, 1989, **3**, 38–39.
11. Griffith, S., *J. Pesticide Reform*, 1989, **9**, 18–20.
12. Firestone, D., *FDA By-Lines*, 1977, No. 2, pp. 57–89.
13. Jensen, A. A., in *Drugs and Human Location* (ed. Bennett, P. N.), Elsevier, Amsterdam, 1988, pp. 551–573.
14. Schecter, A., Startin, J. R., Rose, M., Wright, C., Parker, I., Woods, D. and Hansen, H., *Chemosphere*, 1990, **20**, 919–925.
15. Frommberger, R., *Chemosphere*, 1990, **20**, 333–342.
16. Fries, G. F., *Chemosphere*, 1987, **16**, 2123–2128.
17. McLachlan, M. S., Thoma, H., Reissinger, M. and Hutzinger, O., *Chemosphere*, 1990, **20**, 1013–1021.
18. Bluthgen, A. H. and Heeschen, W. H., Proc. 23rd Internat. Dairy Congress, 1990, vol. 2, pp. 991–1003.
19. Muto, H. and Takizawa, Y., *Arch. Environ. Health*, 1989, **44**, 171–174.

20. Koppe, J. G., Pluim, E. and Olie, K., *J. R. Soc. Med.*, 1989, **82**, 416–419.
21. Vaz, R., Slorach, S. and Albanus, L., *Var Foda*, 1988, **40**, 416–423.
22. Poland, A. and Knudson, J. C., *Annu. Rev. Pharmacol. Toxicol.*, 1982, **22**, 517–554.
23. Poli, A. and Renzoni, G., *Ann. Fac. Med. Vet. Pisa*, 1982, **35**, 81–92.
24. Poli, A. and Renzoni, G., *Ann. Soc. Ital. Sci. Vet.*, 1982, **36**, 529–531.
25. Ryan, J. J., Proc. 23rd Internat. Dairy Congress, 1990, vol. 2, pp. 1387–1392.
26. Heath, R. G., Haress, R. L., Gross, M. L., Lyon, P. A., Dupuy, A. E., Jr. and McDaniel, D. D., *Anal. Chem.*, 1986, **58**, 463–468.
27. Sherry, J., ApSimon, J., Collier, L., Wilkinson, R., Albro, P. and Afghan, B., *Chemosphere*, 1989, **19**, 255–261.

The birth of the semiconductor superlattice*

L. Esaki

IN 1969, research on artificially structured materials was initiated with a proposal for an engineered semiconductor superlattice with a one-dimensional periodic potential by Esaki and Tsu[1,2]. In anticipation of advancement in controlled epitaxy of ultrathin layers, two types of superlattices were envisioned: doping and compositional, as shown at the top and bottom of Figure 1, respectively.

Before arriving at the superlattice concept, we were examining the feasibility of structural formation of potential barriers and wells that were thin enough to exhibit resonant tunneling[3]. A resonant tunnel diode[4,5] appeared to have more spectacular characteristics than the Esaki tunnel diode[6], the first quantum electron device consisting of only a single tunnel barrier. It was thought that advanced technologies with semiconductors might be ready for demonstration of de Broglie electron waves. Resonant tunnelling can be compared to the transmission of an electromagnetic wave through a Fabry–Perot resonator. The equivalent of a Fabry–Perot resonant cavity is formed by the semiconductor potential well sandwiched between the two potential barriers.

The idea of the superlattice occurred to us as a natural extension of double-, triple- and multiple-barrier structures: the superlattice consists of a series of potential wells coupled by resonant tunnelling. An important parameter for the observation of quantum effects in the structure is the phase-coherent, length which approximates to the electron mean free path. This depends on bulk as well as the interface quality of crystals, and also on the temperatures and values of the effective mass. As schematically illustrated in Figure 2, if characteristic dimensions such as superlattice periods or well widths are reduced to less than the phase-coherent length, the entire electron system will enter a mesoscopic quantum regime of low dimensionality, being in a scale between the macroscopic and the microscopic. Our proposal was indeed to explore quantum effects in the mesoscopic regime.

The introduction of the one-dimensional superlattice potential perturbs the band structure of the host materials, yielding a series of narrow subbands and forbidden gaps which arise from the subdivision of the Brillouin

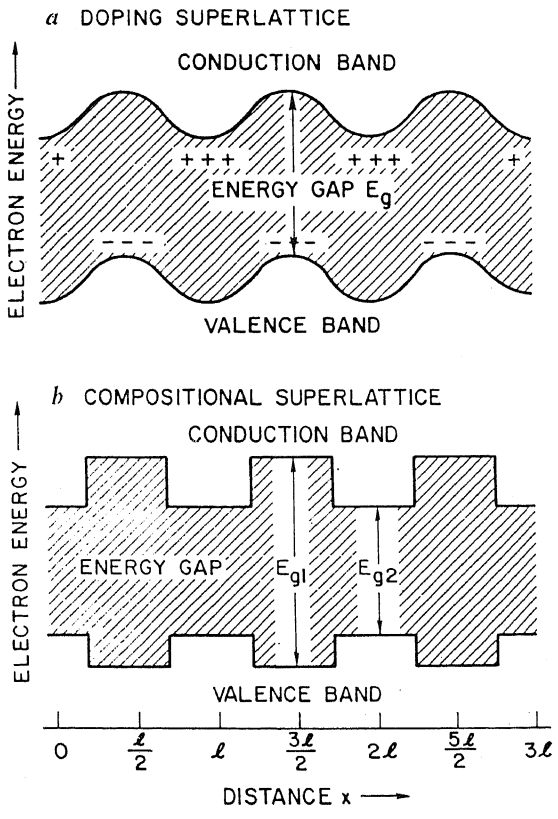


Figure 1. Spatial variation of the conduction and valence band edges in two types of superlattices: *a* doping; *b* compositional.

L. Esaki is the President of the University of Tsukuba, Ibaraki 305, Japan.

*Opening scientific address delivered at the International Conference on Physics and Industrial Development: Bridging the Gap, New Delhi, 17–19 January 1994. This article is reproduced from *Physics and Industrial Development: Bridging the Gap* (Chandrasekhar, S., ed.), Wiley Eastern, 1995. Copies can be procured from the Centre for Liquid Crystal Research, P. O. Box. 1329, Jalahalli, Bangalore 560 013, India.

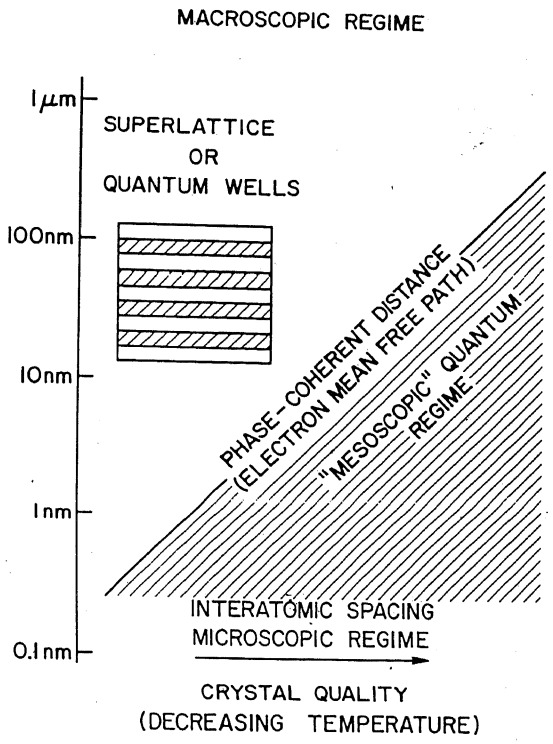


Figure 2. Schematic illustration of a 'mesoscopic' quantum regime (hatched) with a superlattice or quantum wells in the inset.

zone into a series of minizones. Thus, the superlattice was expected to exhibit unprecedented electronic properties. At the inception of the superlattice idea, it was recognized that the long, tailored lattice period provided a unique opportunity to exploit electric-field-induced effects. The electron dynamics in the superlattice direction was analysed for conduction electrons in a narrow subband of a highly perturbed energy–wave-vector relationship. The result led to the prediction of the occurrence of a negative differential resistance at a modestly high electric field, which could be a precursor of the Bloch oscillation. The superlattice, apparently, allows us to enter the regime of electric-field-induced quantization: the formation of Stark ladders,[7,8] for example, can be proved in a (one-dimensional) superlattice,[9] whereas in natural (three-dimensional) crystals the existence and real nature of these localized states in a high electric field have been controversial[10,11].

This was, perhaps, the first proposal which advocated to the engineer with advanced thin-film growth techniques a new semiconductor material designed by applying the principles of the quantum theory. The proposal was indeed made to the US Army Research Office (ARO), a funding agency, in 1969, having daringly stated, with little confidence in a successful outcome at the time, 'the study of superlattices and observations of quantum-mechanical effects on a new physical scale may provide a valuable area of investigation in the field of semiconductors'.

Although this proposal was favourably received by ARO, the original version of the paper[1] was rejected for publication by *Physical Review* on the referee's unimaginative assertion that it was *too speculative* and involved *no new physics*. The shortened version published in *IBM J. Res. Develop.*[2] was selected as a Citation Classic by the Institute for Scientific Information (ISI) in July 1987. Our 1969 proposal was cited as one of the most innovative ideas at the ARO 40th Anniversary Symposium, Durham, North Carolina, 1991.

At any rate, with the proposal, we initiated such a formidable task as to make a *gedanken* experiment a reality. In some circles, the proposal was criticized as close to impossible. One of the objections was that a manmade structure with compositional variations in the order of several nanometers could not be thermodynamically stable because of interdiffusion effects. Fortunately, however, it turned out that interdiffusion was negligible at the temperatures involved.

In 1970, Esaki *et al.*[12] studied a GaAs–$GaAs_{0.5}P_{0.5}$ superlattice with a period of 20 nm synthesized with CVD (chemical vapour deposition) by Blakeslee and Aliotta[13]. Although transport measurements failed to reveal any predicted effect, the specimen probably constituted the first strained-layer superlattice having a relatively large lattice mismatch. Early efforts in our group to obtain epitaxial growth of $Ge_{1-x}Si_x$ and $Cd_{1-x}Hg_xTe$ superlattices were soon abandoned because of the rather serious technological problems at that time. Instead, we focused our research effort on compositional GaAs–$Ga_{1-x}Al_xAs$ superlattices grown with MBE (molecular beam epitaxy). In 1972, Esaki *et al.*[14] found a negative resistance in such superlattices, which was, for the first time, interpreted in terms of the superlattice effect.

Following the derivation of the voltage dependence of resonant tunnel current[5], Chang *et al.*[15] observed the current–voltage characteristics with a negative resistance. Subsequently, Esaki and Chang[16] measured quantum transport properties in a superlattice with a narrow bandwidth, which exhibited an oscillatory behaviour. Tsu *et al.*[17] performed photocurrent measurements on superlattices subject to an electric field perpendicular to the plane layers with the use of a semitransparent Schottky contact, which revealed their miniband configurations.

Heteroepitaxy is of great interest for the growth of compositional superlattices. Innovations and improvements in epitaxial techniques such as MBE and MOCVD (metalloorganic chemical vapour deposition) have made it possible to prepare high-quality heterostructures. Such structures possess predesigned potential profiles and impurity distributions with dimensional control close to interatomic spacing. This great precision has cleared access to the mesoscopic quantum regime[18,19].

Since a one-dimensional potential can be introduced along with the growth direction, famous examples in the

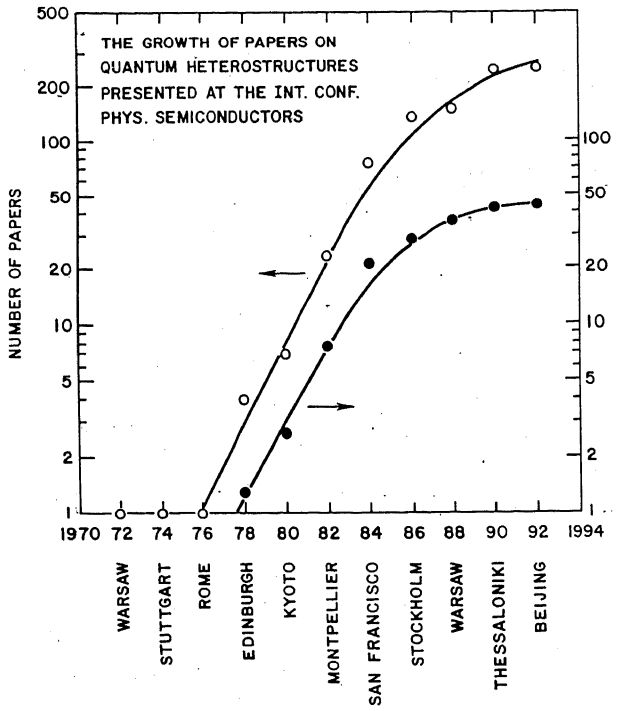


Figure 3. Growth in relevant papers at the biennial International Conference on the Physics of Semiconductors.

history of one-dimensional mathematical physics, including the above-mentioned resonant tunnelling[3], Kronig–Penney bands[20], Tamm surface states[21], Zener band-to-band tunnelling[22], and Stark ladders including Bloch oscillations[7–9], all of which had remained more or less textbook exercises, could, for the first time, be practised in a laboratory. Thus, do-it-yourself quantum mechanics is now possible, since its principles dictate the details of semiconductor structures[23].

We have witnessed remarkable progress in semiconductor research of superlattices and quantum wells over the last two decades. Our original proposal[1] and pioneering experiments apparently triggered a wide spectrum of experimental and theoretical investigations on this subject. A variety of engineered structures exhibited extraordinary transport and optical properties which may not even exist in any natural crystal. Thus, this new degree of freedom offered in semiconductor research through advanced material engineering has inspired many ingenious experiments, resulting in observations of not only predicted effects but also totally unknown phenomena. As a measure of the growth of the field, Figure 3 shows the number of papers related to the subject and the percentage of the total presented at the biennial International Conference on the Physics of Semiconductors. After 1972, when the first paper[14] came on the scene, the field went through a short period of incubation and then experienced a phenomenal expansion in the 1980s. It appears that nearly one-half of semiconductor physicists in the world are working in this area. Activity in this new frontier of semiconductor physics has in turn given immeasurable stimulus to device physics, provoking new ideas for applications. Thus, a new class of transport and optoelectronic devices has emerged. In this interdisciplinary research, there have been numerous beneficial cross-fertilizations. I hope this article provides some flavour of the excitement associated with the birth of the semiconductor superlattice.

1. Esaki, L. and Tsu, R., *IBM Research Note RC-2418*, 1969.
2. Esaki, L. and Tsu, R., *IBM J. Res. Develop.*, 1970, **14**, 61.
3. Bohm, D., *Quantum Theory*, Prentice Hall, Englewood Cliffs, N. J., 1951, p. 283.
4. Iogansen, L. V., *Zh. Eksp. Teor. Fiz.*, 1963, **45**, 207. *Sov. Phys. JETP*, 1964, **18**, 146.
5. Tsu, R. and Esaki, L., *Appl. Phys. Lett.*, 1973, **22**, 562.
6. Esaki, L., *Phys. Rev.*, 1958, **109**, 603.
7. James, H. M., *Phys. Rev.*, 1949, **76**, 1611.
8. Wannier, G. H., *Elements of Solid State Theory*, Cambridge University Press, Cambridge, 1959, p. 190; *Phys. Rev.*, 1960, **117**, 432.
9. Shockley, W., *Phys. Rev. Lett.*, 1972, **28**, 349.
10. Zak, J., *Phys. Rev. Lett.*, 1968, **20**, 1477; *Phys. Rev.*, 1991, **B43**, 4519.
11. Rabinowitch, A. and Zak, J., *Phys. Rev.*, 1971, **B4**, 2358.
12. Esaki, L., Chang, L. L. and Tsu, R., Proc. 12th Int. Conf. Low Temp. Phys., Kyoto, Japan, 1970, p. 551.
13. Blakeslee, A. E. and Aliotta, C. F., *IBM J. Res. Develop.*, 1970, **14**, 686.
14. Esaki, L., Chang, L. L., Howard, W. E. and Rideout, V. L., Proc. 11th Int. Conf. Phys. Semiconductors, Warsaw, Poland, 1972, p. 431.
15. Chang, L. L., Esaki, L. and Tsu, R., *Appl. Phys. Lett.*, 1974, **24**, 593.
16. Esaki, L. and Chang, L. L., *Phys. Rev. Lett.*, 1974, **33**, 495.
17. Tsu, R., Chang, L. L., Sai-Halasz, G. A. and Esaki, L., *Phys. Rev. Lett.*, 1975, **34**, 1509.
18. Esaki, L., *IEEE J. Quant. Electron.*, 1986, **QE-22**, 1611.
19. Esaki, L., in *Highlights in Condensed Matter Physics and Future Prospects* (ed. Esaki, L.), Plenum Press, New York, 1991, p. 55.
20. Kronig, R. de L. and Penney, W. G., *Proc. R. Soc. (London)*, 1931, **A130**, 499.
21. Tamm, I., *Phys. Z. Sowjetunion*, 1932, **1**, 733.
22. Zener, C., *Proc. R. Soc. (London)*, 1934, **145**, 523.
23. Esaki, L., *Phys. Scr.*, 1992, **T42**, 102.

NMR studies on the Myb-DNA system

P. K. Radha, Anup Madan, L. C. Padhy and R. V. Hosur

Tata Institute of Fundamental Research, Homi Bhabha Road, Bombay 400 005, India

NMR structural studies on a bacterially produced DNA binding domain of *Drosophila* c-Myb protein as well as its cognate DNA sequence carried out in our laboratory have been surveyed. The structure of a self-complementary dodecamer DNA containing the Myb responsive element (TAACGG) has been determined to atomic resolution by the combined use of two-dimensional NMR, spectral simulations, restrained energy minimization and distance geometry calculations. The structure is seen to possess novel features which may play important roles during its interaction with the Myb protein. The DNA binding domain of c-Myb protein was seen to have a hydrophobic core and we have identified the types of residues contributing to its formation. Residues contributing to the hydrophobic core formation are seen to be well spread out over the whole length of the 160 residues in the protein and include isoleucines, valines, leucines, alanines threonines, aromatic residues, glutamines and possibly aspartates. Our experimental data in combination with those of others indicate that some of the amino acid residues which form the helical motifs that directly interact with DNA may also be a part of the hydrophobic core.

NUCLEAR magnetic resonance (NMR) spectroscopy has grown leaps and bounds during the last decade, establishing itself as the single most reliable technique for precise determination of three-dimensional (3D) structures of biological macromolecules in solution[1]. Starting from the discovery of Fourier transform NMR[2] in the 1960s, a series of developments such as multidimensional NMR[3], rotating frame experiments[4], inverse detection techniques[5] and gradient enhancement techniques[6] have contributed to the possibility of 3D structure determination of proteins and nucleic acids having molecular weights (MWs) as large as 20 kDa. Nearly a hundred protein structures and a similar number of nucleic acid structures in the MW range of 5–15 kDa have been determined to date with the above techniques. The developments in NMR have been facilitated by parallel developments in other areas, namely computer technology and genetic engineering, the former contributing to data-handling, spectrometer control and data-processing abilities and the latter contributing to the production of desired molecules in large quantities. The developments are still occurring at a rapid pace and the objective is to push the molecular size limits to even higher ranges of 30–40 kDa. These developments will take us in the realm of biology, providing insights into structure–function relationships in biomolecules.

The c-Myb protein is a proto-oncogene product which binds to DNA in a sequence-specific manner[7,8]. The protein functions as a transcriptional regulator in vertebrates and activates a large number of genes corresponding to different proteins[9–16]. The DNA binding activity of the protein is mediated by a region close to its amino terminus[17,18]. This domain contains three imperfect, tandem repeats of 51–53 amino acid residues and is represented in general as R1–R2–R3 (or R1R2R3). Each of the repeat units contains three tryptophan residues that are evenly spaced, 18–19 amino acids apart, and are evolutionarily conserved. Besides, the sequences corresponding to the DNA binding domains of the Myb proteins are highly conserved through evolution and have been detected in as widely divergent species as yeast, *Arabidopsis*, *Zea Mays*, *Drosophila*, chicken, mouse and man[19,20]. It is, therefore, envisioned that the tryptophan residues play an important role, directly or indirectly, in sequence-specific DNA binding. This notion is further supported by site-directed mutagenesis experiments, where replacement of tryptophans at many positions was found to affect drastically the DNA binding properties of the R1R2R3 protein[21].

The three-dimensional structure of the R1R2R3 protein is not yet known. To date, there have been three reports of nuclear magnetic resonance (NMR) studies; one on the solution structure of R3 repeat[22] and two on the R2R3 repeats[23,24] from two different species, mouse and chicken. The sequence homology in these proteins is more than 70%, being much better in the R3 domain. In the mouse protein it was observed that both R2 and R3 repeats contain three helices each and the typical helix–turn–helix motif was present in both the domains. In the case of chicken it was observed that the R3 domain contained three helices with a helix–turn–helix motif, while the R2 domain contained only two helices. The R1 domain exhibited largest variations in sequence homology and its exact role in the DNA binding process is not yet fully understood. It is, however, known that the presence of the R1 domain enhances the DNA binding affinity of the protein.

The consensus nucleotide sequence of Myb recognition has been identified to be YAACKG; more recently, this has been extended to 8 bases as YAACKGHH, where Y = C/T, K = G/T and H = A/C/T[25]. However, since the extension has substantial variability, it is unlikely to be

involved in specific contacts with the protein. There are conflicting reports in the literature as to whether the conformation of the protein changes on complex formation with DNA or not. From gel electrophoresis experiments, indications have come that the interaction induces bending in DNA[26]. Recent NMR results on the R23-DNA complex[24] have shown that the core AAC sequence makes the maximum specific contacts with the protein side-chains and the phosphate interactions provide a greater stability to the complex.

In the last few years we have studied the structure and interactions of the *Drosophila melanogaster* (*Dm*) c-Myb protein with its specific cognate DNA sequences. We have bacterially overexpressed the DNA binding domain of *Dm*-Myb, a polypeptide of 160 amino acid residues, and examined its structure and interactions by various spectroscopic techniques[27–31]. This article gives a survey of the significant achievements, with emphasis on the NMR-derived results. First we describe the solution structure of the Myb responsive element (MRE) determined to atomic resolution, and then describe the structural aspects of the hydrophobic core, which probably plays an important role in spatially orienting the helical segments directly involved in protein–DNA interactions.

Solution structure of the cognate DNA sequence

The DNA sequence considered here is d-ACCGTTAACGGT, the underlined stretch being the Myb recognition sequence. The molecule is self-complementary and, hence, exists as a duplex in solution under the usual conditions of room temperature, neutral pH and low salt, 0.1 M NaCl. The structure has been determined by combined use of 2D NMR, restrained energy minimization and distance geometry calculations, whose details have appeared elsewhere[32]. Briefly, the following steps were involved. First, sequential resonance assignments were obtained by standard procedures based on 2D NMR[1]. Then 2D J-correlated spectra were quantitatively interpreted for extracting coupling constant information and NOE correlated spectra were semiquantitatively interpreted to obtain rough estimates of the interproton distances. In the next stage, quantitative interpretation of the NOESY cross-peaks intensities in terms of 3D structure was carried out by explicit relaxation matrix calculations and iterative comparisons of the calculated and the experimental intensities. This step is described in some detail in the following paragraphs.

Experimental peaks in a 300 ms NOESY spectrum were integrated above suitable thresholds by summing up all the points spanning the peak. Calculated NOE intensities for any starting structure were obtained by relaxation matrix calculation[33,34] followed by a scaling procedure inherent in the SIMNOE algorithm[35]. All the integrated intensities thus obtained were referenced to a particular peak in both experimental and calculated spectra for normalization and comparison. In the absence of a good fit, the structure was refined and the procedure repeated until a satisfactory fit resulted. The iterative structure refinement proceeded as follows. A set of distance constraints with upper and lower bounds was derived by looking at the NOE fits at any particular step. Then restrained energy minimization with the NOE potential defined in terms of distance violations was carried out using the software package X-PLOR[36], for about 20–200 cycles. NOE fits were checked and the distance constraint set redefined to drive the structure towards a better fit. The restrained energy minimization included also the 120 dihedral angle constraints to restrict the sugar geometries in both the strands to the

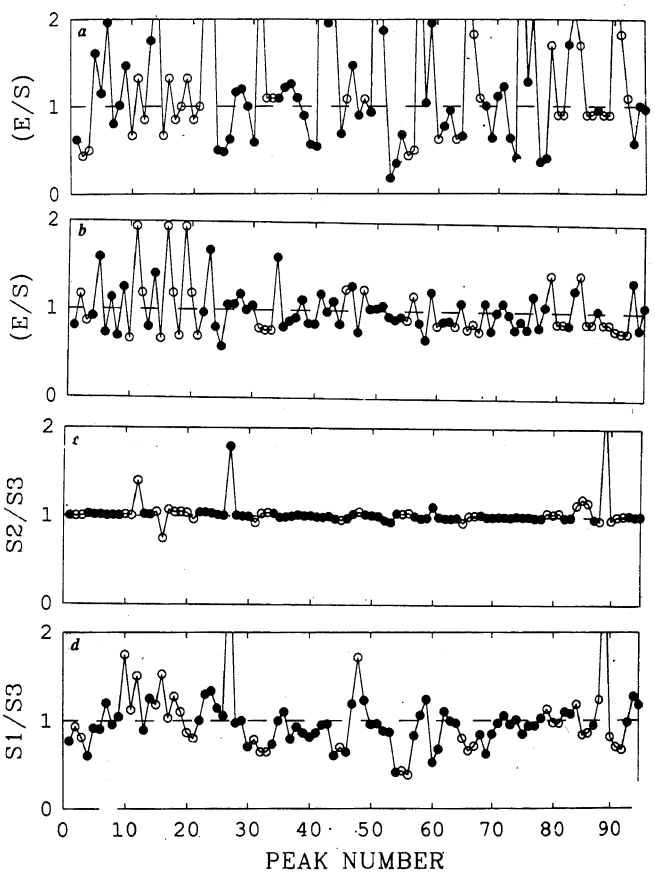


Figure 1. *a*, Comparison of experimental and calculated NOE peak intensities for the initial (B-DNA) model. Open circles represent overlapping peaks and filled circles nonoverlapping peaks. The peaks have been arranged sequencewise from left to right. *b*, Peak-to-peak fit of experimental and calculated NOE peak intensities for the final structure. *c*, Peak-to-peak fit for the two structures S2 and S3 differing by only 0.05 Å on the all-atom r.m.s.d. scale. *d*, Comparison of simulated NOE intensities for the final structure S1 and another distance geometry generated structure S2, which differs from S1 to an extent of 0.68 Å on the all-atom r.m.s.d. scale.

ranges derived from the coupling-constant information. The whole procedure was repeated with different starting models, and we noted that an energy-minimized B-DNA as the starting structure provided the best NOE fits at the end.

Figures 1 *a* and *b* show the initial and final fits for 95 NOE peaks for each strand of the 12-mer duplex. The NOE set consisted of 61 intranucleotide and 34 internucleotide peaks. To appreciate the adequacy of the NOE set and the quality of the fit in terms of closeness of the structure to the correct structure, we have carried out the following exercise. We selected two intermediate structures labelled S2 and S3, which had an all-atom r.m.s.d. (root mean square deviation) of 0.69 and 0.68 Å, respectively, from the final structure (labelled S1). We then compared the calculated NOEs of all the three structures and these are depicted in Figures 1 *c*, *d* in the same manner as the fits in Figure 1 *b*. These results provide a correlation between the similarity of structures and the similarity of NOE fits for the 95 peaks used. The two structures S2 and S3 which have an r.m.s.d. of 0.05 Å and thus are clearly identical have almost identical NOEs (Figure 1 *c*); the differences seen reflect the small differences in structures. Further, the structures S1 and S2, which are also close to the extent of 0.69 Å on the r.m.s.d. scale, have NOEs differing by 25–50% for a large number of peaks; a few peaks differ to a wider extent. This indicates that even as small a structural difference as 0.69 Å on the r.m.s.d. scale results in

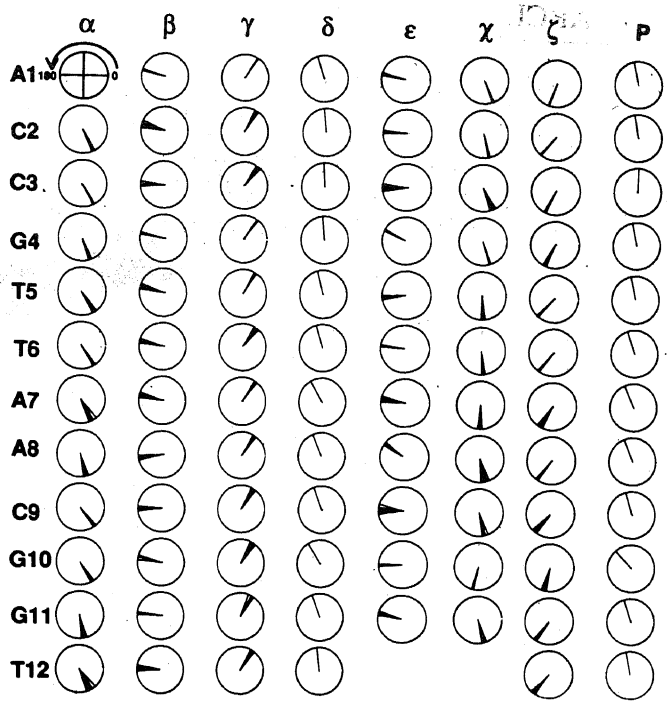


Figure 2. Analysis of the backbone torsion angles of 25 DG-generated structures. The individual torsion angles for all the nucleotides belonging to different structures are indicated by the value in degree in the corresponding dial. The direction of rotation is indicated in the first dial. The torsion angle is indicated in the top row and the nucleotide number in the leftmost column. As can be noted from the analysis, the torsion angles for all the residues are similar within ~ 5–10° for all the structures, indicating a single dominant structure for the DNA segment.

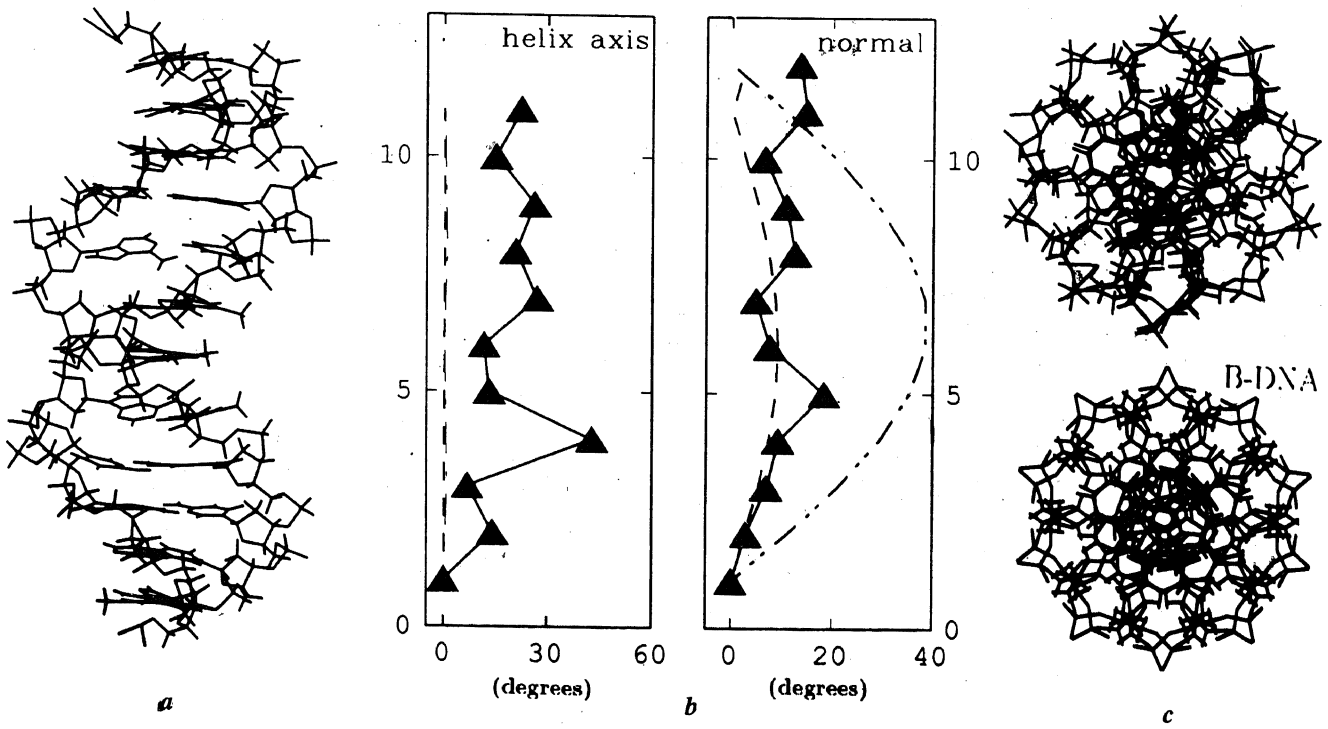


Figure 3. *a*, Stereo view of the final energy-minimized structure. *b*, The course of the local helix axis and the base pair normals with respect to the first base pair as the reference. The standard values for A-DNA (– . .) and B-DNA (. . . .) models have also been indicated. *c*, Comparison of the orthogonal views of the final structure and the B-DNA structure.

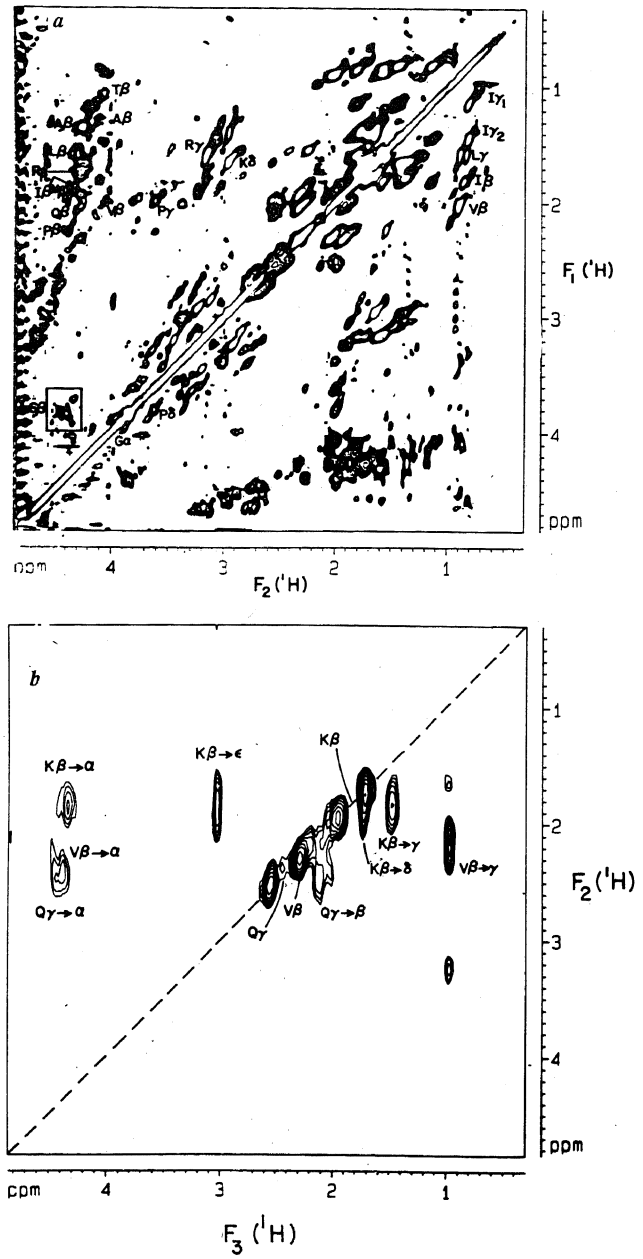


Figure 6. Experimental illustration of the spectral simplifications achieved in the 3D spectrum as compared to the 2D spectrum. *a*. Portion of 2D ^{1}H–^{1}H clean TOCSY spectrum. *b*, A ^{13}C plane through the full 3D spectrum, at a ^{13}C shift of 33.4 ppm. A number of correlations which are hard to find out clearly in the top spectrum are readily identified in the lower spectrum and these have been indicated. The spin systems of K, V and Q residues can be clearly seen.

solvent, and loss of secondary structure elements indicated that the hydrophobic core may be the most stable element since it is the last one to get disrupted by urea. The fluorescence results also indicated that all the three repeats, R1, R2 and R3, in the protein contribute to the formation of the hydrophobic core. Because of these significant observations, we have tried to characterize the nature of the hydrophobic core in terms of identifying the participating amino acids by NMR spectroscopy.

Specific monitoring and distinction of the interior residues from the exterior ones can be made in NMR by making use of the fact that the amide and other side-chain NH_2 groups lying on the surface of the protein exchange protons more rapidly with solvent compared to those in the interior of the protein. Deuterium exchange studies provide the most direct method to identify the slowly exchanging amides (exchange rates $< 10^3\ s^{-1}$) in the protein. Amides exchanging faster than this rate will not be visible in the spectra in ^2H_2O solutions. In the case of R1R2R3 protein, we observed that only four or five out of the 162 NH–C$^\alpha$H cross-peaks could be counted in the clean TOCSY spectrum in ^2H_2O solution at 16°C and these were present even after two months of keeping the protein in ^2H_2O. This implied that a small fraction of the protein is deeply buried in the hydrophobic environment and the rest of the protein is easily accessible to the solvent. We resolved to attempt a finer gradation of the exchange rates of amides by performing temperature variation experiments in H_2O solutions, that is, if the amide proton resonances are monitored as a function of temperature, the fast exchanging amides disappear from the spectra faster and this provides a useful filter to study the slowly exchanging amide protons. We may mention here that the ^{1}H spectra in ^2H_2O solutions did not show perceptible major changes as the temperature was raised slowly from 16 to 80°C. This indicated rather qualitatively that the structural core of the protein did not undergo major changes, and also, more relevantly in the present context, that the chemical shift assignments we would obtain at a higher temperature can be extrapolated to lower temperatures.

Figure 5 shows the same regions of clean TOCSY spectra of R1R2R3 protein at 16 and 80°C recorded with the same mixing time of 80 ms. Several points are noteworthy from these two spectra: (i) All the side-chain NH_2's of lysines and arginines, which can be easily identified from their characteristic relay peaks to other nonexchangeable protons along their chains, have exchanged out with water at 80°C. This implies that these two residue types lie entirely on the surface of the protein. (ii) Quite substantial simplification of the spectrum has occurred at 80°C and it is possible to count peaks in different regions; nearly 75 NH–C$^\alpha$H peaks could be counted in the fingerprint region of the spectrum at 80°C. Here it may be noted that several C$^\alpha$H protons occurring at 4.20–4.25 ppm are killed by water presaturation at 80°C. (iii) Several NH–C$^\alpha$H–C$^\beta$H relay peaks arising from aromatic residues and possibly aspartates were detectable. This identification is provisional and is based on the typical chemical shifts of C$^\beta$H protons. Actually, C$^\beta$H protons of aspargines,

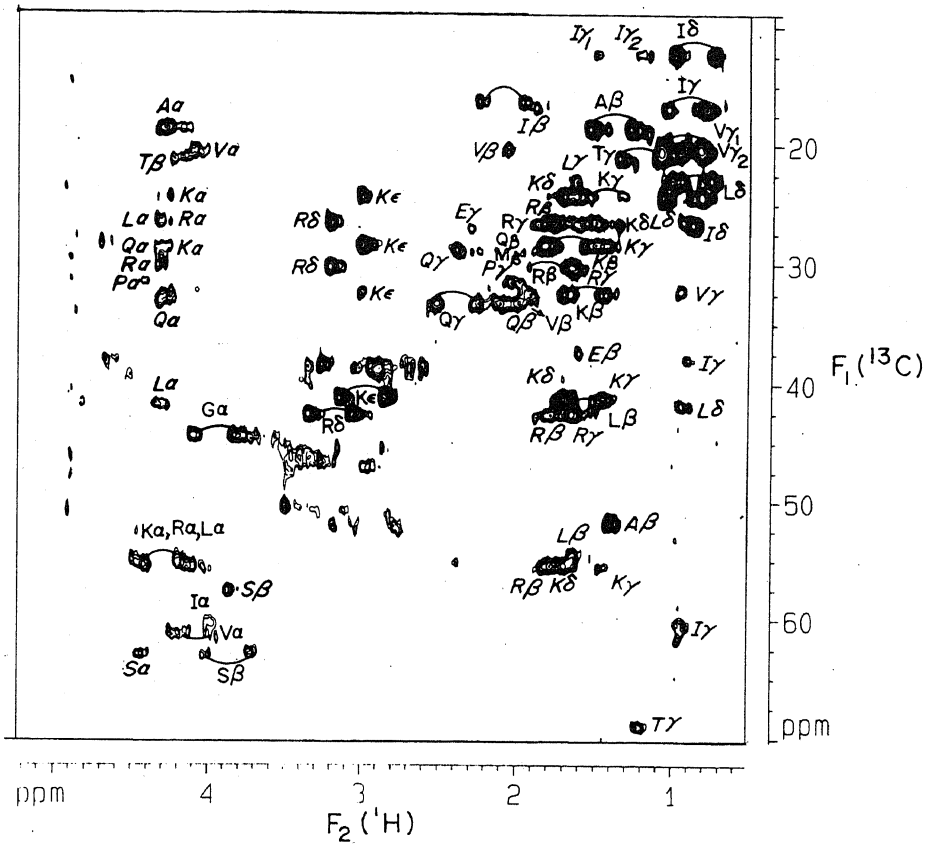


Figure 7. ^{1}H–^{13}C HMQC-TOCSY spectrum at natural abundance of the R1R2R3 protein recorded on BRUKER AMX 500 NMR spectrometer. Sample conditions: 5 mM protein, pH 5.6, temperature 16°C. 2048 t_2 and 512 t_1 points were used with carbon offset placed in the middle of the C^{β} region. The carbon spectral width was 90 ppm. Quadrature direction along F_1 was achieved by the standard TPPI procedure. The MLEV mixing time was 51 ms. No ^{13}C decoupling was employed during detection and thus the direct peaks in the spectrum retain the one-bond $^{1}J_{CH}$ coupling (they have been joined). The relay peaks are, however, single peaks. The various peak assignments indicated have been obtained by analysis of different 2D spectra and the 3D spectrum.

glutamines and cysteines are also expected in this region, but the presence of aspargines and glutamines can be easily checked from the double quantum (DQ) spectra. Indeed, the DQ spectra in H_2O showed that aspargines were absent but some glutamines were present whose $C^{\beta}H$ positions have consequently been identified. The presence or absence of cysteine was not possible to ascertain from these experiments. The number of remaining peaks suggested that both aromatic and aspartate residues were retained. (iv) There are clearly several methyl-bearing residues such as isoleucines, leucines, alanines, threonines and valines, and the NH protons of these show a reasonable spread.

Additional residue type identification of peaks at 80°C was obtained from the heteronuclear ^{1}H-^{13}C correlated 2D and 3D spectra. The spectra shown in Figure 6 illustrate how 3D spectra can be used for spin system identifications. The spectrum in (*a*) is the usual clean TOCSY spectrum in $^{2}H_2O$, which contains all the observable ^{1}H-^{1}H correlations and relay peaks for the protein. Clearly, the spectrum is highly crowded and difficult to analyse for unambiguous connectivities. The spectrum in (*b*) is a slice through the three-dimensional ^{13}C separated ^{1}H–^{1}H clean TOCSY spectrum taken at a particular carbon chemical shift. The spectrum is now simpler and the peaks can be readily identified as originating from particular spin systems. Such an analysis of all possible ^{13}C planes allows the identification of all the amino acid residues in the protein. Figure 7 summarizes all the identified spin systems in a 2D HMQC-TOCSY spectrum, which is basically a ^{1}H–^{13}C correlation spectrum including the ^{1}H–^{1}H relays at specific ^{13}C chemical shifts. The peaks in this spectrum provide a cross-check for the validity of the assignments made from 3D spectral analysis. An important observation made from this spectrum is that the methyl-bearing residues which are difficult to distinguish in ^{1}H–^{1}H correlated spectra are easily distinguished on the basis of the ^{13}C chemical shifts.

The above assignments provide a basis for the identification of the peaks in Figure 5. Figure 8 shows the specific side-chain assignments to residue types obser-

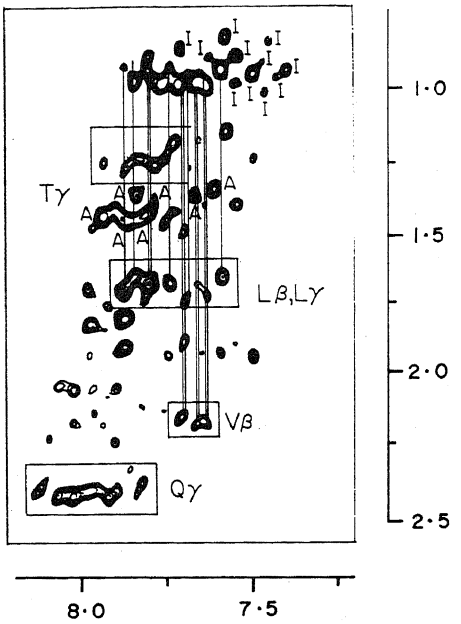


Figure 8 Specific side-chain assignments to residue types observed in the clean TOCSY spectrum (80 ms spin lock time) of R1R2R3 protein in H_2O recorded at 80°C. Cross-peaks belonging to valines and leucines are indicated by vertical lines. Peaks belonging to glutamine and threonine γ have been boxed, and those belonging to alanines have been labelled A. The 11 isoleucine γ's have been labelled as I.

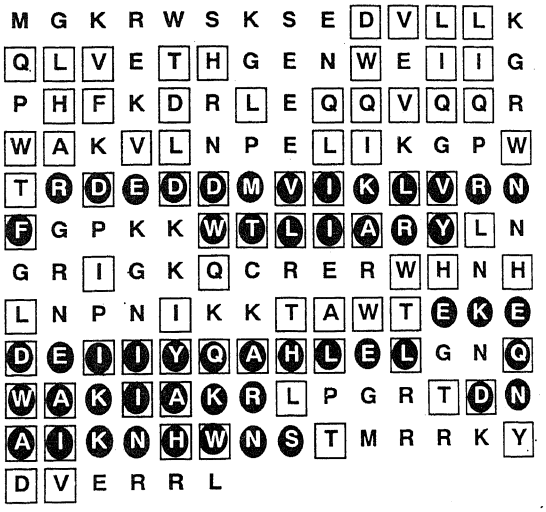


Figure 9. Amino acid sequence of R1R2R3 protein indicating the residues (boxed) for which the amides are observed at 80°C. All the residues (excepting the terminal ones) of any particular type that has been seen in the high-temperature spectrum have been marked, although we were able to count the numbers only in few cases. For example, separate identification of all peaks belonging to individual aromatic residues was not done. But the total count in the figure tallies closely with the total number of residues seen in the spectrum. We notice that the slowly exchanging amides are spread out all along the sequence, implying participation of all the portions of the protein in the formation of the hydrophobic domain. The circled residues are expected to be forming helices, from sequence comparison with R2–R3 domains of mouse and chicken c-Myb, whose NMR secondary structures have been reported recently[23, 24].

vable at 80°C that probably constitute the hydrophobic core in the R1R2R3 protein. We noticed here that backbone amide protons had reasonable spread and we could count 5–6 valines, 7 alanines, 6 threonines, 5–6 leucines, 11 isoleucines and 7–8 glutamines. There were several other peaks which originated, very likely, from aspartates and aromatic residues; the latter have hydrophobic character and hence are expected to remain shielded from the water.

Figure 9 shows the sequence-wise location of the various residues identified in Figure 8. We notice that the residues are distributed along the whole length of the protein. This indicates that all the three repeats contribute to the formation of the hydrophobic core in the protein and about 70–80 residues contribute to its formation. It is also worthwhile noting that glutamines and aspartates, which are hydrophilic residues, are also included in the above list. A possible explanation could be that they are involved in the formation of hydrogen bonds through their polar side-chains with other residues in the protein. Examination of the relative positions of the hydrophobic residues in the sequence, *vis-à-vis* the known secondary structure elements of vertebrate R1R2R3 from NMR studies in the conserved R3 and R2–R3 domains and their DNA complexes[22–24], reveals that some of the hydrophobic core residues may form parts of helices involved in specific interactions with the DNA. The helices may be lost at elevated temperatures but the hydrophobic core appears to be the last one to get disrupted[31].

Conclusions and future directions

We have described in this article structural information derived from NMR studies in our laboratory, on a specific target DNA sequence of the DNA binding domain of *Drosophila melanogaster* c-Myb protein and also on the hydrophobic core present in the protein where the DNA binds. The structure determination has unravelled important features which could constitute hot points of recognition. Extensive use of homonuclear and heteronuclear two- and three-dimensional NMR experiments has enabled gross identification of residue types involved in the formation of the hydrophobic core in the protein, which has an important role to play in DNA recognition. The hydrophobic core seems to be constituted by all the three repeats in the protein. Sequence comparison with R2–R3 domain of mouse and chicken c-Myb, for which NMR structures were reported recently, provided clues to identification of the secondary structure and of the so-called helix–turn–helix motif of DNA recognition. This provides a supplementary character to our previous work on R1R2R3–DNA interaction by fluorescence quenching studies, in that the cognate

DNA binding site of this protein is associated with a hydrophobic core element being surrounded by 4–5 tryptophans, at least one tyrosine and a cysteine[27–29].

The work reported in this paper represents an important step forward in our efforts to understand DNA–Myb interaction and recognition at the atomic level. A lot more remains to be done from the viewpoint of 3D structure determination of the protein and of the complex. Sequence-specific assignment of resonances is the most crucial step requiring the employment of a battery of multidimensional NMR experiments on isotopically (^{15}N and ^{13}C) labelled protein. Indeed, we have just successfully achieved the production of large amounts of labelled R1R2R3 protein.

Atomic-level structure determinations of the R1R2R3 and also of their segments R1R2, R2R3 and R3, which are also in the offing in our laboratory are expected to throw light on several important questions. For example: (1) What is the role of R1 repeat in specific recognition? (2) Do the individual repeats retain their structural motifs when any of the other repeats is knocked off from the protein? (3) Which are the residues making contact with the DNA and how is the specificity of interaction brought about? (4) Is it possible to alter the DNA binding specificity by structural-knowledge-based mutations in the protein?

1. Wuthrich, K., *NMR of proteins and Nucleic Acids*, Wiley, New York, 1986.
2. Ernst, R. R. and Anderson, W. A., *Rev. Sci. Instrum.*, 1966, **37**, 93–102.
3. Ernst, R. R., Bodenhausen, G., and Wokaun, A., *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*, Clarendon Press, Oxford, 1987.
4. Bothner-By, A. A., Stephens, R. L., Lee, J., Warren, C. D. and Jeantoz, R. W., *J. Am. Chem. Soc.*, 1984, **106**, 811–813.
5. Muller, L., *J. Am. Chem. Soc.*, 1979, **101**, 4481–4484.
6. Davis, A. L., Keeler, J., Laue, E. D. and Moskau, D., *J. Magn. Reson.*, 1992, **98**, 207–216.
7. Biekenkapp, H., Borgmeyer, U., Sippel, A. E. and Kempnauer, K. H., *Nature*, 1988, **335**, 835–837.
8. Nakagoshi, H., Nagase, T., Kanei-Ishii, C., Ueno, Y., and Ishii, S., *J. Bio. Chem.*, 1990, **265**, 3479–3483.
9. Ness, S. A., Maarknell, A. and Graf, T., *Cell*, 1989, **59**, 1115–1125.
10 Dasgupta, P., Saikumar, P., Reddy, C. D. and Reddy, E. P., *Proc. Natl. Acad. Sci. USA*, 1990, **87**, 8090–8094.
11. Dasgupta P., Reddy, C. D., Saikumar, P. and Reddy, E. P., 1980, *J. Virol.*, **66**, 270–276.
12. Evans, J. L., Moore, T. L., Kuehl, W. M., Bender, T. and Ling, J., *Mol. Cell. Biol.*, 1980, **10**, 5747–5752.
13. Nicolaides, N. C., Gualdi, R., Casadevall, C., Manzella, L. and Calabretta, B., *Mol. Cell. Biol.*, **11**, 6166–6176.
14. Zobel, A., Kalkbrenner, F., Guemann, S., Nawarath, M., Vorbruggen, G. and Moelling, K., *Oncogene*, 1991, **6**, 1397–1407.
15. Siu, G., Wurster, A. L., Lipsick, J. S. and Hedrick S. M., *Mol. Cell. Biol.*, 1992, **12**, 1592–1604.
16. Sureau, A., Soret, J., Vellard, M., Crochet, J. and Perbal, B., *Proc. Natl. Acad. Sci. USA*, 1992, **89**, 11683–11687.
17. Katzen, A. L., Kornberg, T. B. and Bishop, J. M. *Cell*, **41**, 449–456.
18. Peters, C. W. B., Sippel, A. E., Vlngron, M. and Klempnauer, K.-H., *EMBO J.*, 1987, **6**, 3085–3090.
19. Luscher, B. and Eiseman, R., *Genes Develop.*, 1990, **4**, 2235–2240.
20. Shen-Ong, G. L. C., *Biochem. Biophys. Acta*, 1990, **1032**, 39–52.
21. Saikumar, P., Murali, R. and Reddy, E. P., *Proc. Natl. Acad. Sci. USA*, 1990, **87**, 8452–8456.
22. Ogata, K., Hojo, H., Aimoto, S., Nakai, T., Nakamura, A. S., Ishii, S. and Nishimura, Y., *Proc. Natl. Acad. Sci. USA*, 1992, **89**, 6428–6432.
23. Jamin, N., Gabrielsen, O. S., Gilles, N., Lirsac, P.-N. and Toma, F., *Eur. J. Biochem.*, 1993, **216**, 147–154.
24. Ogata, K., Morikawa, S., Nakamura, H., Sakikawa, A., Inoue, T., Kanai, H., Sarai, A., Ishii, S. and Nishimura, Y., *Cell*, 1994, **79**, 639–659.
25. Weston, K. *Nucleic Acids Res.*, 1992, **20**, 3043–3049.
26. Saikumar, P., Gabreil, J. L. and Reddy, P., *Oncogene*, 1994, **9**, 1279–1287.
27. Madan, A., Hosur, R. V. and Padhy, L. C., *Biochemistry*, 1994, **33**, 7120–7126.
28. Radha, P. K., Madan, A., Padhy, L. C. and Hosur, R. V., *Curr. Sci.*, 1994, **66**, 287–294.
29. Madan, A., Radha, P. K., Hosur, R. V. and Padhy, L. C., *Eur. J. Biochem.*, 1995, in press.
30 Radha, P. K., Madan, A., Padhy, L. C. and Hosur, R. V., *Proc. Indian. Acad. Sci. Chem. Sci.*, 1994, **106**, 7, 1537–1549.
31. Madan, A., Radha, P. K., Srivastava, A., Hosur, R. V. and Padhy, L. C., *Eur. J. Biochem.*, 1995, in press.
32. Radha, P. K., Madan, A., Nibedita, R. and Hosur, R. V., *Biochemistry*, 1995, **34**, 5193–5922.
33. Macura, S. and Ernst, R. R., *Mol. Phys.*, 1980, **41**, 95–117.
34. Keepers, J. W., and James, T., *J. Mag. Reson.*, 1984, **57**, 404–426.
35. Nibedita, R., Ajay Kumar, R., Majumdar, A. and Hosur, R. V., *J. Biomol. NMR*, 1992, **2**, 467–476.
36 Brunger, A. T., X-PLOR (Version 2.1) Manual, Fellows of Harvard University, New York, 1990.
37. Radha, P. K., Nibedita, R., Ajay Kumar, R. and Hosur, R. V., *Methods Enzymol.*, **261**, in press.

ACKNOWLEDGEMENTS. The facilities provided by the National Facility for High-Field NMR supported by the Department of Science and Technology are gratefully acknowledged. We thank Bhabha Atomic Research Centre for the use of X-PLOR package and Dr Manju Bansal for the NUPARM package for DNA structure analysis.

Lithium and tungsten mineralization in Sewariya pluton, South Delhi Fold Belt, Rajasthan: Evidences for preferential host rock affinity

Subrata Banerji and M. K. Pandit*

B-130, Malviya Nagar, Jaipur 302 017, India
*Department of Geology, University of Rajasthan, Jaipur 302 004, India

'Sewariya batholith', the largest of Erinpura granite intrusions in the South Delhi Fold Belt is evolved from a two-stage subduction process. The Sewariya granitoids comprise three major mineralogical varieties, viz. hornblende–biotite granite, biotite granite and leucogranite. The hornblende–biotite granite is a deep-level I-type granite, whereas the other two are orogenic, S-type anatectic granites formed by partial melting of the same source material under varying physicochemical conditions. These two granite types also host the tungsten and lithium mineralization. The W and Li minerals show perfect separation and host rock discrimination. The W minerals are associated with biotite granite, whereas the leucogranite hosts the Li mineralization.

THE folded rock sequences of northwestern part of the Indian shield in the state of Rajasthan contain two principal well-defined Proterozoic fold belts, the early Proterozoic Aravalli and the Middle to Late Proterozoic Delhi Fold Belt. The Delhi Belt extending from Delhi in the northeast to northern Gujarat in the southwest is exposed in two discrete sectors[1], viz. North Delhi Fold Belt (NDFB) in northeastern Rajasthan and South Delhi Fold Belt (SDFB) towards southwest, with its best development in the main Aravalli hill range near Ajmer and in western Mewar. The Delhi sequence has been traced into southwestern Rajasthan but its physical continuity is interrupted by a vast expanse of alluvium, resulting in tremendous correlation uncertainties. Orogenic activities manifested by continuous polyphase acid magmatism have been documented in NDFB by 1600–1500 Ma granites and by 850–700 Ma plutonic belt in SDFB, overlapped and followed by the Malani igneous activity (740–430 Ma). The SDFB is intruded by a number of granitoid bodies ranging from minor intrusions to composite batholiths. The 22 mile long and 4 mile wide Sewariya batholith is the largest of the Erinpura granite intrusions in the northern half of the Delhi synclinorium[1].

Location and geological setting

The present work deals with the evidences of typical mutual association of specific granitoid host *vis-à-vis* tungsten and lithium mineralization in the area south of Luni river in northwestern flank of SDFB (26°30″ to 26°32′N and 74°05′ to 74°22′E) around village Sewariya in district Pali, Rajasthan (Figure 1). The geological setting of the area is shown in the detailed lithologic map (Figure 1) and the stratigraphic sequence is summarized in Table 1.

The oldest rocks encountered in the area are the Pre-Delhi metamorphites comprising mica schist, gneisses and amphibolites. These are overlain by two thick bands of marble, commonly known as 'Ras Marble'. The contact between Ras Marble and Pre-Delhi rocks is either concealed by alluvium or infested by pegmatites. The Ras Marble is coarse-grained, light grey and thickly bedded massive carbonate rock containing minor quantities of quartz, felspar along with epidote, zoisite, diopside and tremolite as accessories. A volcano-sedimentary assemblage constituting the Barotiya Group containing meta-argillite (+/– autoclastic conglomerate), tourmalinite, meta-basic volcanics, meta-gabbro, meta-arkose, marble and chert overlies the Ras sequence. The contact between Barotiya group of rocks and the Pre-Delhi metamorphites/Ras Marble is obliterated and occupied by Sewariya granitoids (SG). The meta-argillite contains quartz, biotite, plagioclase, garnet, tourmaline, staurolite, andalusite, sillimanite (fibrolite), apatite and magnetite. Tourmalinite contains mainly tourmaline and quartz besides minor opaques and biotite as accessories. The meta-volcanics of the Barotiya group exhibit bimodal composition[2]. In the present area only the mafic component having actinolite, plagioclase (andesine), quartz, sodic hornblende, epidote, apatite, leucoxene and opaques is exposed. Relict pyroxene cleavage is observed in the amphiboles. Alteration of hornblende into chlorite and plagioclase into saussurite is also seen. Three petrographic variants of meta-gabbroid bodies have been identified in the field. Marble contains discrete clasts of anhedral quartz in fine micritic mass along with zoisite, clinozoisite and apatite as common accessory minerals. Meta-arkose is dirty-white-coloured, medium-grained rock containing mainly quartz and felspars besides muscovite and apatite along with minor opaques. Meta-chert occurs as thin linear bands within meta-arkose unit in the area.

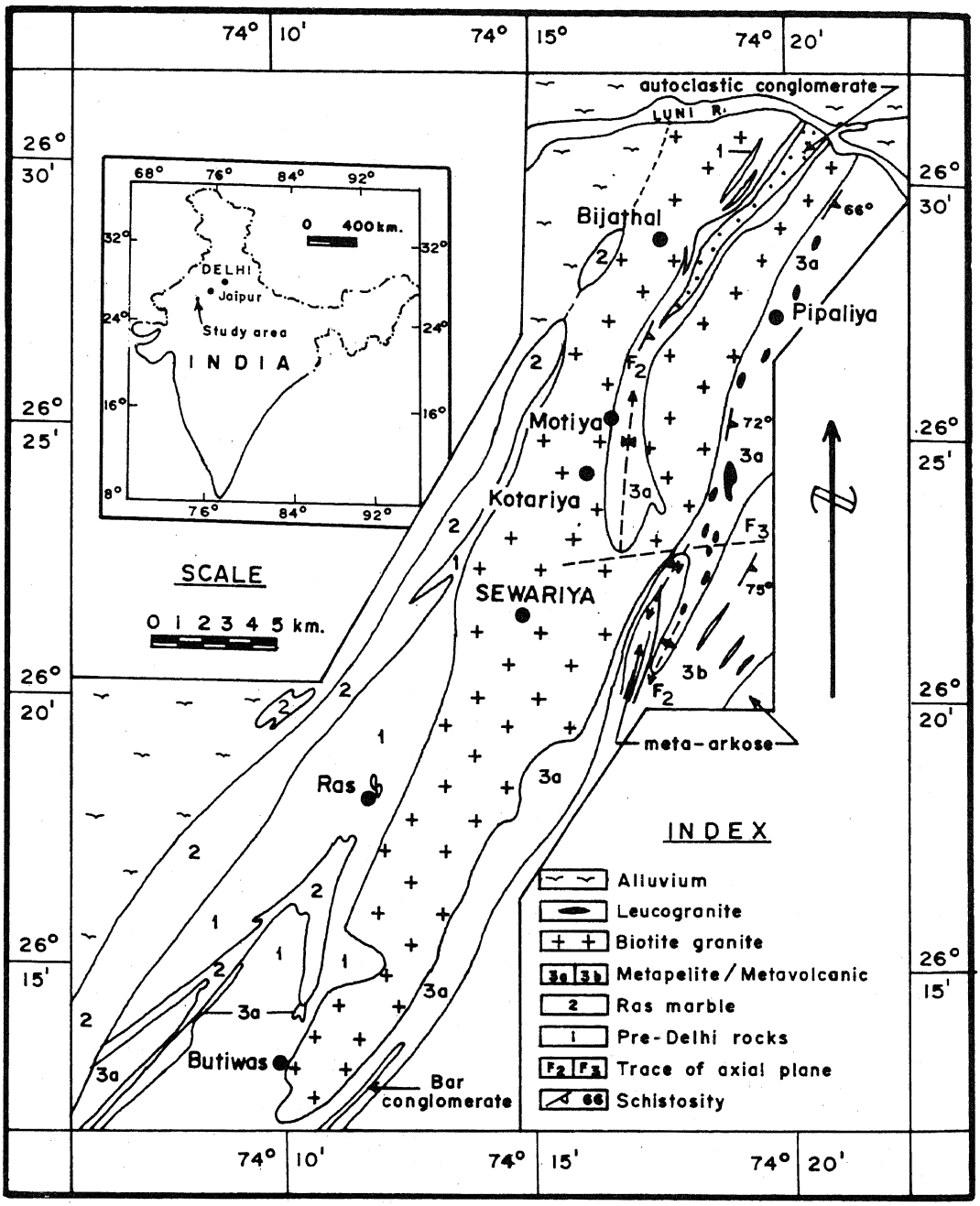


Figure 1. Lithostratigraphic map of the Sewariya region. Map of India showing the location of the study area is also given in the inset.

Various structural elements recorded in the area are in harmony with the deformations related to SDFB. The overall strike pattern, litho-contacts and linear shears of the area are in conformity with NNE–SSW-trending Aravalli orographic axis. Three generations of conformable deformational structures have been recognized in both Ras and Barotiya group of rocks, while one more phase of earlier folding is also recorded in the Pre-Delhi rocks. Shears usually follow the zones of weakness and contribute prominent linearity to the overall outcrop pattern.

The area of the present study has undergone two progressive metamorphic phases involving lower greenschist to sillimanite–almandine subfacies of amphibolite facies and subsequent anatexis, followed by a retrogressive phase of metamorphism under lower greenschist facies. The grade of metamorphism in this sector increases westwards and the isograds are parallel to the regional planar fabric.

Table 1. Stratigraphic succession of the Sewariya region

Syn- to Post-Delhi intrusives		Pegmatites; quartz tourmaline; quartz andalusite; quartz felspar and quartz veins
		Leucogranite; biotite granite; hornblende–biotite granite
Intrusive contact		
Delhi super group	Barotiya group	Marble; meta-chert; meta-arkose; tourmalinite; meta-volcanics; meta gabbro; mica schist (+/- autoclastic conglomerate)
	Ras sequence	Marble
Structural hiatus		
Pre-Delhi		Metamorphites

Granitoids and metallogeny

Three types of granitoids, viz. leucogranite, biotite granite and hornblende–biotite granite, are exposed in the present area. The latter two granite types and their associated W +/– Li +/– Sn mineralization have already been studied in detail[3]. During the present study the authors have identified the leucogranite variety and have also successfully discriminated exclusive association of tungsten mineralization with biotite granite and lithium with leucogranite–pegmatoid phase. The hornblende–biotite granite seems to be absolutely devoid of any mineralization.

Leucogranite is a fine- to medium-grained, white- to dirty-white-coloured rock occurring as small pods and irregular bodies of extremely variable size, in the eastern part of the area within meta- argillites. The smaller bodies and pegmatite veins lie concordant to the planar fabric and are often boudinaged. The typical mineral assemblage of this rock is quartz, plagioclase (mainly albite), K-felspar, muscovite, tourmaline and garnet with small amounts of apatite. Tourmaline occurs either as irregular masses or as euhedral prisms with 10–15% modal abundance. The leucogranite does not contain a single flake of biotite.

The biotite granite and the leucogranite are meta- to per-aluminous, highly fractionated in character, showing depletion of compatible elements like Fe, Mg, Ti, Ca, Ni, etc. (Table 2). The relatively more fractionated (high SiO_2, very low MgO and CaO and higher alkalies), homogeneous leucogranite has characteristic absence of biotite and zircon, implying efficient separation of melt from solid residual phases[4]. The low Li content of leucogranite (65 ppm) can be attributed to the partitioning of Li into the residual melt phase during fractionation and its ultimate segregation as Li phosphate minerals in the late-stage pegmatoid phase associated with leucogranite. The low W content of leucogranite is due to the absence of biotite phase, which is the carrier and the most suitable locale for W concentration.

Table 2. Major and trace element analysis of Sewariya granitoids. The major oxides are represented as wt% and the trace elements in ppm

	Range	Average[1]	Average[2]
SiO_2	67.71–80.18	71.83	73.62
TiO_2	0.07–0.50	0.22	< 0.05
Al_2O_3	9.67–15.60	13.13	14.88
Fe_2O_3	0.40–4.70	1.93	0.76
FeO	1.08–3.17	1.80	0.90
MnO	0.01–0.10	0.04	0.04
MgO	0.05–1.3	0.64	0.04
CaO	0.56–2.48	1.36	0.41
Na_2O	0.14–4.64	2.57	4.86
K_2O	3.72–6.05	4.91	3.53
P_2O_5	0.03–0.22	0.12	0.22
Trace elements			
Ni	10–20	13	< 10
Zr	20–50	28	< 20
Li	20–900	135	65
W	20–100	50	< 50

[1]Average of 15 biotite granites.
[2]Average of 2 leucogranites.
< ≡ Below detection limit.

The biotite granite is heterogeneous in nature, with composition comparable to the metasediments of the area[4]. It is well differentiated, per-aluminous, orogenic, S-type granite with calc-alkaline affinities[3]. The initial I_{Sr} of 0.74552 indicates its crustal derivation from metasedimentary source rock[5]. The higher concentration of W is due to the presence of biotite as an important constituent. Higher values of Li as compared to the leucogranite (host for Li mineralization) is a result of regular partitioning of Li in the crystal phases during fractionation, resulting in its depletion in the residual melt. As such, the quartz and pegmatite veins associated with the biotite granite are devoid of any Li mineralization.

The crustal anatexis and recycling of pre-existent crustal material plays a dominant role in granitoid petrogenesis. Consequent to progressive regional metamorphism under sillimanite–almandine subfacies of amphibolite facies, anatexis takes place, giving rise to different types of granitic magmas depending mainly upon the composition of the source material, its formation level difference and the percentage of partial melting. In this context the leucogranite appears to have been formed at higher crustal levels, generated by comparatively lower percentage of partial fusion of supracrustals[6] with simultaneous removal of melt from the melting zone under water-rich conditions as evident by the presence of extensive pegmatite veins. The melt initially segregated into small bodies, which subsequently united together to form larger bodies and intruded the overlying rocks in the eastern part of the study area.

The SDFB with an ensialic rifting history contains profuse Post-Delhi granites, the older Sendra granite

and the younger Sewariya granite developed in response to the subduction process[7]. A perfect matching of the geological setting and lithological characters of the present area with pyrenes[8,9] indicates that localized rifting events associated with strike–slip motion in SDFB have affected the basin[10]. Deep circulation of ground waters also influenced the anatectic process by controlling the solidus–liquidus temperature and promoted higher percentage of partial melting to produce such a huge biotite granite pluton (Sewariya batholith), the 'largest of Erinpura granite intrusions'. The biotite granite of the area appears to have been derived from the same source material at relatively deeper levels compared to the leucogranite, involving higher percentage of partial melting under essentially 'wet' melting conditions[6].

Both leucogranite and biotite granite are high-level, syn-metamorphic and syn- to post-kinematic, anatectic granites, but the third variety, i.e. the nonfoliated hornblende–biotite granite is post-metamorphic 'deep-level' granite possibly formed either exclusively from deeper crustal material or by mixing of mantle-derived mafic magma and assimilated crustal lithologies.

The western Indian craton and its fold belts have provided important clues to the understanding of different metallotects for various mineral commodities. Incidence of wolframite was first recorded within Sewariya pluton by Rajasthan State Tungsten Development Corporation in 1987 and later a 14 km long tungsten mineralization belt was identified between Kalni in the north and Kotariya in the south[11]. The W +/– Sn +/– Li metallogeny associated with metallogenetically specialized biotite granite has been compared with similar metallogenesis of the Cordilleran arc system[3]. However, the present study reveals that the tungsten metallogeny in the eastern part of the area associated with meta-argillites (viz. Pipaliya and Pipaliya village prospects) seems to be exhalative strata-bound type. Subsequent enrichment in W content during progressive metamorphism is evidenced by enhanced W values in NNE–SSW- trending dewatering quartz veins emplaced along axial plane schistosity (S2). Evidences of later remobilization and enrichment in quartz veins associated with NNE–SSW-trending shear zones near and along the contact between Barotiya group of rocks and anatectic biotite granite have also been observed. In contrast, at Motiya, Kotariya, Bijathal, etc., the host quartz and pegmatite veins are associated with granitization process emplaced in and through channelways in the shear zones. Tungsten mineralization accompanied by greissenization and as placer deposits in Motiya block, has also been identified. The lithium phosphate minerals (sicklerite, ferrisicklerite and triphylite) bearing zoned pegmatoid and quartz veins associated with leucogranite phase occur concordantly within meta-argillites in the eastern part of the present area as an 18 km long belt. The younger discordant veins show remobilization and enrichment. Similar phosphate species of sicklerite–triphylite also occur as important accessory minerals in numerous granitic pegmatites in the Pala area[12]. The phosphate mineralogy and alteration history of these minerals indicate their crystallization at a relatively shallow depth[13]. The crystallization was facilitated by the presence of a coexisting, exsolved, volatile-rich fluid within the pegmatite magma in the upper intermediate zone only[14]. The phosphate–silicate liquid immiscibility is not essential for lithium phosphate formation under water-saturated conditions[15].

Discussion and conclusion

The relationship between leucogranite and the biotite granite in the present area is not clear as the former is occupying the synformal cores and the latter is emplaced in antiformal cores of the same deformation (DF2). Leucogranite does not contain biotite as well as tungsten, whereas biotite is the main mafic mineral in tungsten-bearing biotite granite. Pegmatitic phase is ubiquitous in leucogranite and contains lithium phosphate minerals. Though a few pegmatite veins are associated with biotite granite, they are devoid of any lithium phosphate mineralization. Pneumatolytic quartz veins are a very common feature of biotite granite, which hosts the bulk of the tungsten mineralization.

The absence of biotite and certain elements in well-differentiated leucogranite implies efficient selective separation of melt from solid residual phases during anatexis. The tungsten normally partitions along with biotite and hornblende in a granitic melt[16], consequently, the biotite-free leucogranite of the present area is devoid of any associated tungsten. The absence of absolute pegmatitic conditions and essential exsolved volatile-rich fluid phase in biotite granite have constrained it to host the lithium mineralization. The present area is characterized by three phases of Erinpura granitic activity. The leucogranite and the biotite granite are high-level, syn-metamorphic and syn- to post- kinematic, S-type granite generated by anatexis of Barotiya group supracrustals at different formation levels involving varying percentages of partial melting. Though supracrustals of the area were the source for both leucogranite and biotite granite, effective selective separation during anatexis has caused lithium phosphate mineralization in extremely well-differentiated leucogranite under absolute pegmatitic conditions along with essential coexisting exsolved volatile-rich fluid phase. For want of such suitable conditions the biotite granite could host only the tungsten mineralization. In Stewart pegmatite, only the upper intermediate zone contains Li mineralization, while simultaneously formed lower intermediate zone is devoid of mineralization due to the absence of exsolved fluid phase, essential for mineralization[15]. These obser-

vations also warn about uncertain chances of Li mineralization at deeper levels in the present area. The hornblende–biotite granite in this area seems to be devoid of any mineralization; however, further studies are underway to explore the possibility of its contribution in the metallogeny of the area.

1. Heron, A. M., *Mem. Geol. Surv. India*, 1953, **79**, 1–139.
2. Bhattacharjee, J., Golani, P. R. and Reddy, A. B., *Indian J. Geol.*, 1988, **60**, 191–199.
3. Bhattacharjee, J., Fareeduddin and Jain, S. S., *J. Geol. Soc. India.*, 1993, **42**, 3–16.
4. Banerji, Subrata, Ph D thesis, University of Rajasthan, Jaipur, 1994.
5. Sarkar, G., Bishul, P. K., Guha, D. B., Jain, S. S., Banerjee, S., Bhushan, S. K., Chittora, V. K. and Bhattacharjee, J., *Rec. Geol. Surv. India*, 1993, **126(2)**, 31–33.
6. Wickham, S. M., *J. Geol. Soc. London*, 1987, **144**, 281–297.
7. Sinha Roy, S., *Mem. Geol. Soc. India*, 1988, **7**, 97–107.
8. Wickham, S. M., *J. Petrol.*, 1987, **28**, 127–169.
9. Wickham, S. M. and Oxburgh, E. R., *Tectonophysics*, 1986, **129**, 53–69.
10. Sinha Roy, S., *Indian Mineral.*, 1989, **43**, 231–240.
11. Jain, S. S. and Bhattacharjee, J., *Indian Mineral.*, 1992, **46**, 159–164.
12. Shigley, J. E. and Brown, C. E. Jr., *Am. Miner.*, 1985, **70**, 395–408.
13. Taylor, B. E., Fourd, E. E. and Friederichsen, H., *Contrib. Mineral. Petrol.*, 1979, **68**, 187–205.
14. Jahns, R. H. and Burnham, C. W., *Econ. Geol.*, 1969, **64**, 863–864.
15. Shigley, J. E. and Brown, C. E. Jr., *Am. Mineral.*, 1986, **71**, 356–366.
16. Dakate, Y. G., IGCP 26 Workshop, 1979, Nagpur, India, pp. 25–30.

ACKNOWLEDGEMENTS. SB gratefully acknowledges kind permission of the Director General, Geological Survey of India for this research work for his Ph D. Discussions with Shri V. D. Chande, Deputy Director General and Dr A. R. Das, Director, G.S.I. have been very useful. We thank the anonymous reviewer for his comments on the earlier version of the manuscript. We are indebted to the Director, Rajasthan University Computer Centre for extending facilities.

Received 15 November 1994; revised accepted 19 June 1995

MEETINGS/SYMPOSIA/SEMINARS

ICPIC: Indian Conference on Pattern Recognition, Image Processing and Computer Vision

Place: IIT Kharagpur
Date: 13–15 December 1995

Topics include: Low level processing, Feature extraction, Data clustering applications, Pattern classification, Image segmentation, Parallel vision algorithms and architectures, Object recognition, Multidimensional signal processing, Colour image analysis, Document image analysis, Image restoration, Image interpretation, 3D scene analysis, Shape from shading, Stereo & depth estimation, Dynamic scene analysis, Neural network for pattern recognition and computer vision, Industrial applications, Remote sensing, Active vision, Asic for image processing, Fuzzy sets in image analysis.

Contact: Dr B. N. Chatterji
Programme Chair, ICPIC
Dean, Academic Affairs
IIT Kharagpur 721 302, India
Phone: (03222)-2221 to 2224
Fax: (03222)-2303 and 2605
e.mail: bnc@iitkgp.ernet.in

National Symposium on Frontiers in Applied Environmental Microbiology

Place: School of Environmental Studies
Cochin University of Science and Technology
Cochin
Date: 11–13 December 1995

Topics include: Microbial management of solid and liquid waste, Microbial quality and quality control of potable and recreational waters, Microbial control of insect pests, Biofertilizers, Microbial leaching and biomining, Microbial fouling and corrosion, Pesticide microbiology, Petroleum microbiology, Microbial productivity and diversity in nature and Microbial technology for controlled aquaculture environments.

Contact: Dr I. S. Bright Singh
Convener
National Symposium
School of Environmental Studies
Cochin University of Science and Technology
Fine Arts Avenue, Cochin 682 016, India

The Triple Helix—University–Industry–Government Relations: A Laboratory for Knowledge based Economic Development

Place: Amsterdam
Date: 4–6 January 1996

A more elaborate theme paper is available on request.

Contact: Mr. Henry Etzkowitz
Department of Computer Science
Columbia University
New York 100027, USA.
Tel: (+1)212-939 7028
Fax: 666 0140
e.mail: etz@cs.columbia.edu
or
Mr. Loet Leydesdorff
Department of Science & Technology Dynamics
Nieuwe, Achtergracht 166
1018 WV Amsterdam
The Netherlands
Tel: (+31) 20-525 6598
Fax: 525 6579
e.mail: loet@sara.nl

Chemotherapy reduces human anti-filarial IgE response

M. K. Beuria* and Manoj K. Das†

Parasite Immunology Division, Regional Medical Research Centre (ICMR), Nandankanan Road, Bhubaneswar 751 016, India
*Present address: Department of Immunology, Central JALMA Institute of Leprosy (ICMR), Agra 282 001, India

The effect of diethylcarbamazine treatment on levels of IgE antibodies to infective larval (L_3) antigen of *Wuchereria bancrofti* and to a purified allergenic fraction (Sd_{30}) of *Setaria digitata* was followed up to 1 year in asymptomatic microfilaraemic patients infected with *W. bancrofti*. IgE levels increased transiently by day 15, followed by marked decline beginning at one month and persisting up to 12 month post-treatment. IgE responses to Sd_{30} underwent greater changes compared to L_3 antigen of *W. bancrofti*.

FILARIAL infection leads to an enhanced production of IgE antibodies in humans. The serum levels of specific IgE in different categories of *Wuchereria bancrofti* infection have been determined[1,2]. The highest level is exhibited by patients with chronic infection, followed by clinically normal people residing in the endemic regions and asymptomatic microfilaraemic carriers (AS). Immunologic changes in microfilaraemic (MF) patients associated with treatment with diethylcarbamazine (DEC), the drug of choice in filariasis, have been studied[3,4]. Chemotherapy brought about enhanced cellular immune responsiveness[5]. Changes in the antibody isotypes, especially IgE, have not received adequate attention. We report here the effect of DEC treatment in MF patients on levels of IgE antibodies to a purified allergenic fraction of *Setaria digitata* adult worms and to somatic infective larval (L_3) antigen of *W. bancrofti*. The allergenic fraction, which has been shown to cross-react serologically with L_3 antigen, exhibits IgE profile in filarial patients similar to that of L_3 antigen[6].

The study population was from a *W. bancrofti* endemic region of Orissa (Khurda district), India[2]. Eighteen microfilaraemic individuals of both sexes, aged 19–46 years (median 25 years), wishing to participate in chemotherapy were given DEC orally for 12 days at 6 mg/kg body weight under the supervision of a local medical officer. Blood samples were collected 15 days, 1, 3 and 8 months and 1 year after treatment. All the individuals were MF-negative when checked at 1 month post-treatment. L_3 antigens of *W. bancrofti* and the allergenic fraction of *S. digitata* (designated as Sd_{30}, molecular weight 30 kDa as determined by SDS-PAGE and G-200 Sephadex Gel filtration) were prepared as described earlier[2,6]. Serum IgE levels (at 1/100-fold dilution) were determined by ELISA by coating the plates with high antigen concentrations. The procedure was followed as described earlier[2].

All patients had measurable levels of IgE antibodies to both antigens before treatment. Changes in levels of IgE antibodies after DEC treatment relative to pretreatment levels are shown in Figure 1 and Table 1. A significant increase ($P < 0.05$) was noticed by day 15 in IgE response to Sd_{30} antigen. This was followed by a sharp decline, leading to values considerably lower than pretreatment values. The reduction observed at 1 month reached a nadir around 3 months after treatment and then increased progressively. The enhancement on day 15 was noticed in all patients except in four (two unaltered; two decreased). At 3 months, 15 patients exhibited a reduction in IgE response to Sd_{30} antigen, ranging from 72 to 100%; only two patients showed 50% and 20% drop. One patient did not exhibit any change after therapy. Levels of IgG antibodies to these antigens were, however, not affected (data not shown) by DEC treatment. All patients were found to remain amicrofilaraemic at 1 year post-treatment. It is of interest to note that IgE levels registered an enhancement at 1 year from the lowest values obtained at 3 months post-treatment.

Levels of IgE antibodies to Sd_{30} allergen underwent

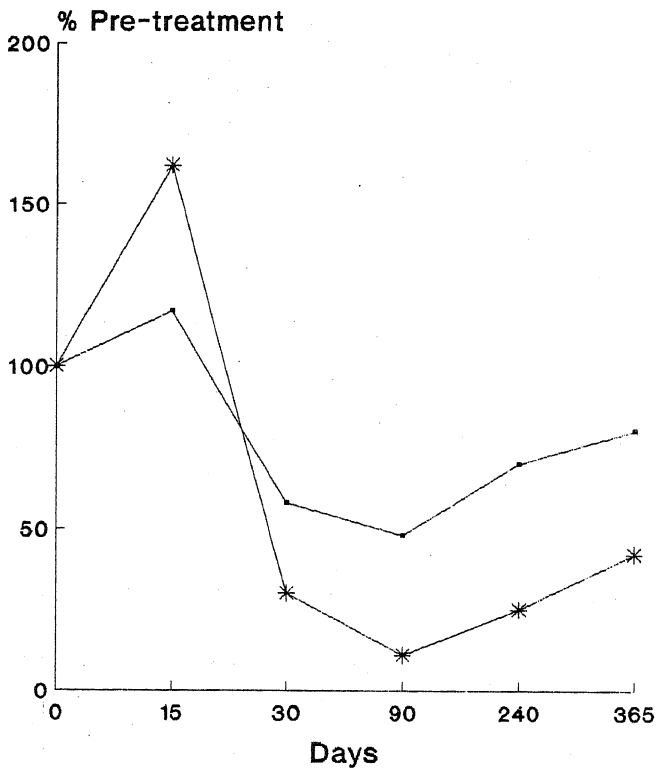


Figure 1. Antifilarial IgE levels after DEC treatment expressed as percentages of pretreatment for each patient. Each point represents the mean value of change with respect to pretreatment response for individual patients. (*) Sd_{30}; (o) L_3 antigen.

†For correspondence.

Table 1. IgE levels (ELISA $A_{492} \pm SD$) in microfilaraemic patients ($n = 18$) after diethylcarbamazine treatment

	A_{492}					
	Pretreatment	Days posttreatment				
Antigen	Day 0	15	30	90	240	365
Sd_{30}	0.22 ± 0.10	0.36 ± 0.16 **	0.06 ±0.04 *	0.02 ± 0.02 *	0.05 ± 0.04 *	0.08 ± 0.06 *
W. bancrofti L_3	0.34 ± 0.10	0.40 ± 0.22 NS	0.19 ± 0.11 *	0.16 ±0.10 *	0.24 ± 0.13 *	0.30 ± 0.18 NS

$^{**}P < 0.05$ compared to day 0; $^{*}P < 0.01$ compared to corresponding day 0 or day 15.
NS: not significant compared to day 0.

greater changes than those to *W. bancrofti* L_3 antigen. For example, the mean reductions in IgE response to Sd_{30} and L_3 antigen 3 months after treatment were 90% and 50% of pretreatment values, respectively. The lower reduction in L_3-IgE may be caused by persistent exposure to *W. bancrofti* larvae in the endemic region. It might also be possible that DEC treatment induces changes in IgE response to epitope(s) selectively present in Sd_{30} allergen. Earlier we have demonstrated that IgE response to these antigens is filariae-specific since people living in nonfilarial regions of Orissa exhibited negligible IgE levels[2,6]. IgE production is sustained in people in endemic areas by them being continuously exposed to the infection since clinically normal individuals of the endemic region have high IgE levels[2].

Considering the biologic roles of IgE antibodies in parasite infections, which vary from allergic disorders to conferring protection to hosts[7,8], the effect of drug treatment on IgE response is an interesting topic. The changes in specific IgE levels as reported here have not been characterized in human filariasis. The present report describes the response of the levels of IgE antibodies to two different antigens, L_3 stage of *W. bancrofti*, which initiates the infection, and a purified filarial allergen, Sd_{30}. Both the antigens exhibited diminished IgE production in treated patients over a course of time.

Immunologic changes following a two-year community trial of ivermectin treatment in Guatemala of patients infected with *Onchocerca volvulus* have been reported[9]. Enhanced parasite-specific IgE levels were noticed 6 months after treatment, which fell marginally below the pretreatment level by the second year. Although antiallergic action of DEC is well documented[10–12], the underlying mechanism was not known. The present results demonstrate clearly that DEC treatment diminishes the production of filarial IgE antibodies in humans and thereby contributes to decreasing allergic reactions.

The long-term monitoring of the study population will be helpful in finding out when IgE levels would be generated again in the DEC-treated patients.

1. Hussain, R., Hamilton, R. G., Kumarswami, V., Franklin, A. N. and Ottesen, Jr. E. A., *J. Immunol.*, 1987, **127**, 1623–1628.
2. Das, M. K., Beuria, M. K. and Dash, A. P, *Int. Arch. Allergy Immunol.*, 1992, **99**, 118–122.
3. Lammie, P. J., Eberhard, M. L., Lowrie, R. C. and Katz, S. P., *Trans. R. Soc. Trop. Med. Hyg.*, 1988, **82**, 726–729.
4. Ramprasad, P., Prasad, G. B. K. S. and Harinath, B. C., *Acta Tropica*, 1988, **45**, 245–255.
5. Piessens, W. F., Ratiwayanto, S., Pissens, P. W., Tuti, S., Mc Greevy, P. B., Darwis, F., Palmieri, J. R., Koiman, I. and Dennis, D. T., *Acta Tropica*, 1981, **38**, 227–234.
6. Beuria, M. K. and Das, M. K., *J. Biosci.*, 1992, **17**, 435–461.
7. Oglivie, B., *Nature*, 1964, **204**, 91–92.
8. Hagan, P., *Parasite Immunol.*, 1993, **18**, 1–4.
9. Steel, K., Lujan-Trangay, A., Gonzalez-Peralta, C., Zea-Flores, G. and Nutman, T. B., *J. Infect. Dis.*, 1991, **164**, 581–587.
10. Thiruvengadam, K. V., Subramaniam, N., Devarajan, T. V and Zachariah, M. G. M., *J. Indian Med. Assoc.*, 1974, **63**, 278–281.
11. Murthy, P. K., Katiyar, J. C., Chandra, R., George, P. A. and Sen, A. B., *Indian J. Med. Res.*. 1978, **68**, 428–434.
12. Mackenzie, C. B., *Trop. Dis. Bull.*, 1985, **82**, R1–R37.

ACKNOWLEDGMENT. We thank the Director General, ICMR for support.

Received 24 March 1995; revised accepted 20 June 1995

Biochemical basis for the differentiation of the two nonpoisonous snakes *Eryx conicus* Schneider and *Eryx johnii* Russell (family: Boidae)

P. Senthil Kumar, A. P. Kamalakara Rao* and T. M. Vatsala

Shri AMM Murugappa Chettiar Research Centre, Tharamani, Madras 600 113, India
*Pachaiyappa's College, Madras 600 013, India

The tissue-specific patterns of lactate dehydrogenase (LDH) for two species of the genus *Eryx* are reported and characterized utilizing sodium deoxycholate. LDH-1 is predominant in *Eryx conicus*, whereas LDH-3 and LDH-2 are observed in *Eryx johnii*. Sodium deoxycholate is a selective inhibitor of LDH-5 in *Eryx johnii*. Variations are observed in all the tissues between these two species and there is a variability of tissue LDH expressions within the species.

ISOENZYMES are multiple molecular forms of enzymes[1]. They can serve in taxonomic, genetic, phylogenetic and

evolutionary studies and help to solve many biological problems[2–5]. Lactate dehydrogenase (lactate: NAD oxido reductase, EC 1.1.1.27) controls the interconversion of pyruvate and lactate in the glycolytic pathway and thereby serves as an important source of the oxidized coenzyme (NAD) during periods of transient anaerobiosis. In most of the vertebrates, the isoenzymes of LDH are formed by the random self-assembly, *in vivo* and *in vitro*, of two different polypeptides A and B to form all possible isozymes, the two homopolymers A_4 and B_4 and the three heteropolymers A_3B_1, A_2B_2 and A_1B_3. The specificity of LDH polypeptide association is brought about by a combination of noncovalent forces[6,7].

LDH constitutes an excellent model system for the analysis of the problems of polypeptide assembly. Depending upon the organism and tissue, LDH can exhibit self-assembly[8,9], epigenetic modification of assembly[10–12], genetically controlled restriction of assembly[13–17] and directed assembly.

In this paper we report our investigation of the electrophoretic characteristics of LDH from two species of nonpoisonous snakes belonging to the genera *Eryx* of the family Boidae. Our aim is to determine the tissue-specificity of LDH subunit synthesis, to look for variations at the LDH locus and to compare isoenzyme patterns from the tissues of these two species.

Two species of snakes, namely *Eryx conicus* and *Eryx johnii*, were used in this study of the LDH tissue patterns. The snakes, all adults, used in this investigation were collected from different localities of Madras.

Twelve tissues, namely eye lens, brain, trachea, lungs, heart, liver, stomach, intestine, muscle, kidney, vertebrae and skin, were dissected from fresh specimen. Care was taken to see that the eye lens was free from contamination by the vitreous humour, retinal and choroid elements. Brain, lungs, heart and liver were washed thoroughly with ice-cold phosphate buffer of pH 7.4 as to render them free from traces of blood. Stomach and intestine were washed till they were free from their contents. Other tissues were washed thoroughly with phosphate buffer of pH 7.4. All the chemicals used in this study were purchased from E. Merck (West Germany).

The tissues were homogenized in iced condition with phosphate buffer of pH 7.4. The homogenates were centrifuged at 4°C at 25,000 × *g* for 30 min. The supernatants were used for gel electrophoresis.

A 5.5% polyacrylamide gel was used[18]. The current flow was adjusted to 2.5 mA per sample and run for 45 min with Tris–glycine buffer of pH 8.3. The gels were removed and stained. Inhibition studies were carried out using sodium deoxycholate.

The main aim was to determine the tissue-specificity of the LDH subunit synthesis in order to pinpoint the species specific biochemical markers and also to compare the LDH isoenzyme patterns in these two species.

The LDH patterns of the 12 tissues of *Eryx conicus* and *Eryx johnii* were observed and are shown in Figures 1 and 2, respectively.

The LDH isoenzymes of the snakes studied exhibited a net negative charge difference between isoenzymes. Therefore, the isoenzymes migrated more slowly towards the anode at pH 8.3.

In *Eryx conicus*, of the 12 tissues, 10 showed the predominance of only LDH-1 (B_4) isoenzyme – vertebrae and skin exhibited the four-isoenzyme pattern, LDH-5 (A_4), LDH-4 (A_3B_1), LDH-3 (A_2B_2) and LDH-2 (A_1B_3).

In *Eryx johnii*, except for lungs and vertebrae, all the other tissues analysed exhibited two isoenzyme fractions LDH-3 (A_2B_2) and LDH-2 (A_3B_1). Lungs and vertebrae showed the presence of three isoenzymes, namely LDH-4 (A_3B_1), LDH-3 (A_2B_2) and LDH-2 (A_1B_3).

The Rf values of these LDH isoenzymes were 0.18, 0.24, 0.28, 0.32 and 0.34 for LDH-5, LDH-4, LDH-3, LDH-2 and LDH-1, respectively, and are presented in Figure 3.

There was widespread resemblance among the LDH isoenzymes of these two species and a characteristic pattern of electrophoretic mobilities of LDH-5 (A_4), LDH-4 (A_3B_1), LDH-3 (A_2B_2), LDH-2 (A_1B_3) and LDH-1 (B_4). No difference was found in the LDH patterns of males and females. Each tissue exhibited a characteristic LDH pattern and the tissue-specific difference in the relative staining intensities of each LDH isoenzyme attributable to the quantitative variation in the proportion of the isoenzymes, which has been a function of a particular ratio of A and B subunits of each tissue. The staining intensity is shown in the densitometric scanning of the LDH isoenzyme of the two species, illustrated in Figures 4 and 5.

Sodium deoxycholate was used as a selective inhibitor of LDH-5 (A_4) isoenzymes to establish the subunit composition of LDH isoenzymes and to determine their genetic basis. There was a selective inhibition of A isoenzymes. The B subunits were resistant to the action of sodium deoxycholate. The characterization of the LDH isoenzymes was performed based on the literature[19].

The importance of this study lies in highlighting as to what extent the LDH isoenzymes play a role in the establishment of the basic tenets of biochemical taxonomy.

The differential regulation of gene function is a fundamental aspect of cellular differentiation. Isoenzyme (multiple molecular forms of enzymes) are ideal gene products for the analysis of differential gene function during embryogenesis and cellular differentiation[20–22].

Differences in the tissue pattern of enzyme locus expression were used to infer the extent of divergence among species. Some locus tissue expressions diverged markedly among species, while others were conserved.

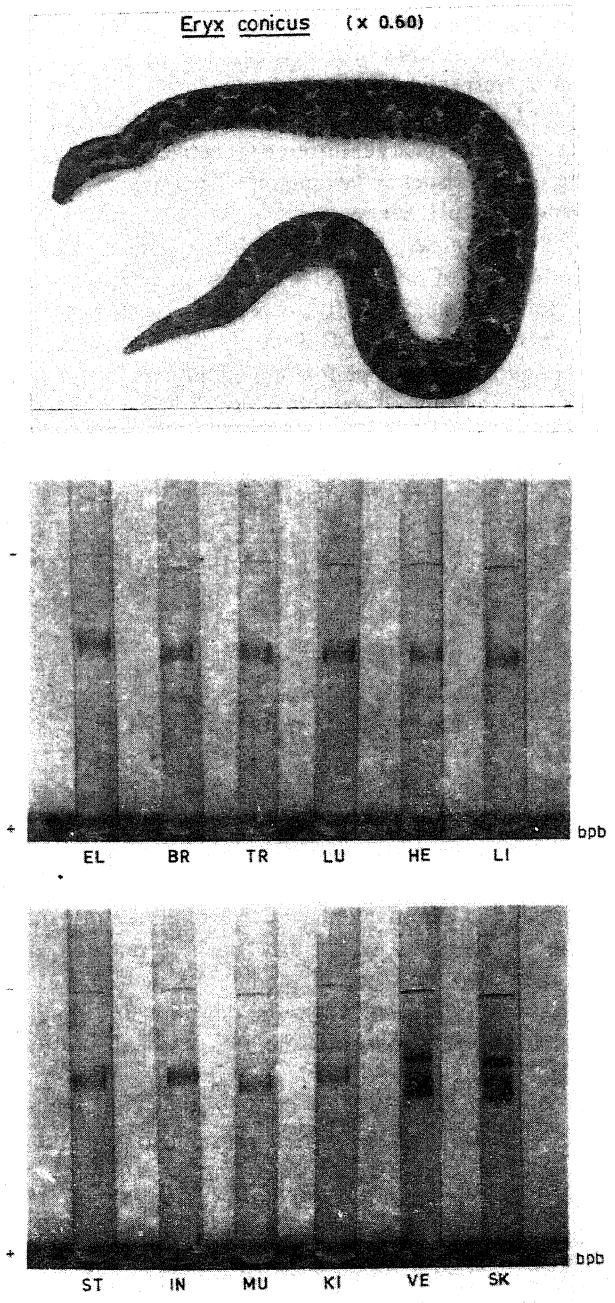


Figure 1. LDH isoenzyme profile of twelve tissues of *Eryx conicus*.

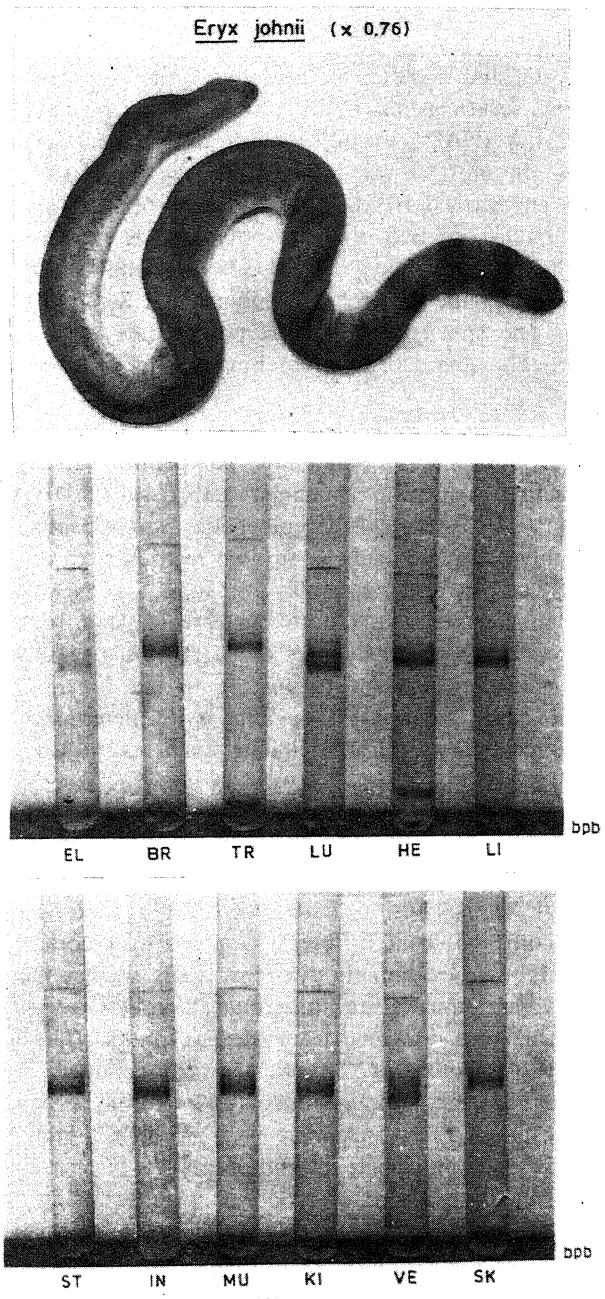


Figure 2. LDH isoenzyme profile of twelve tissues of *Eryx johnii*.

Differences in activity for a given enzyme among tissues, as well as for different enzymes within the same tissue, were sufficiently independent to permit each locus tissue expression of a species to be treated as a separate character. Statistically significant difference in levels of tissue enzyme activities among species were then used to construct a phylogeny. The phylogeny constructed using tissue enzyme expressional differences was similar to that based on enzyme structural difference and to one of morphologically based phylogenies[23].

The partial functional redundancy of L-lactate dehydrogenase (LDH) (EC 1.1.1.27) isoenzymes has permitted marked difference in the tissue regulations of their genes[24–29]. Both differential synthesis and catabolism contribute to the tissue patterns of LDH gene expression[30–34]. Tissue differences in the steady-state levels of LDH isozymes are probably due to mutations at sequences other than those coding for the respective isoenzymes. Such mutations could affect the number of enzyme molecules by altering one or more processes, ranging from transcriptions to catabolism, that is, by altering the regulations of gene expression in the broadest sense.

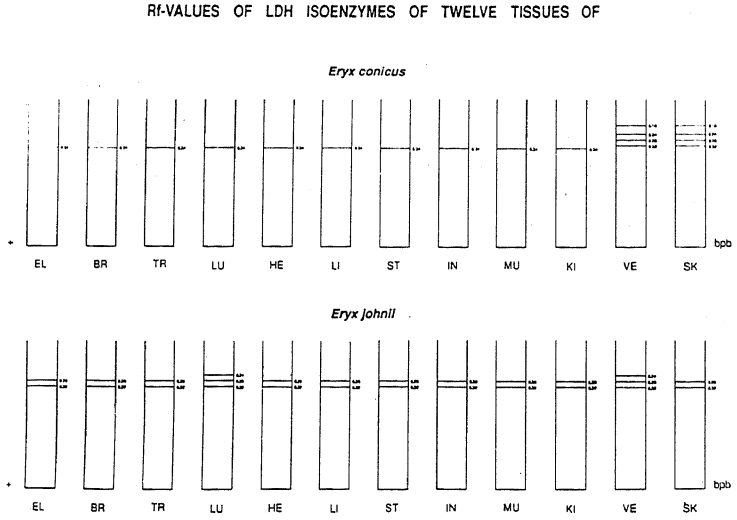


Figure 3. Rf-values of LDH isoenzymes of twelve tissues of *Eryx conicus* and *Eryx johnii*.

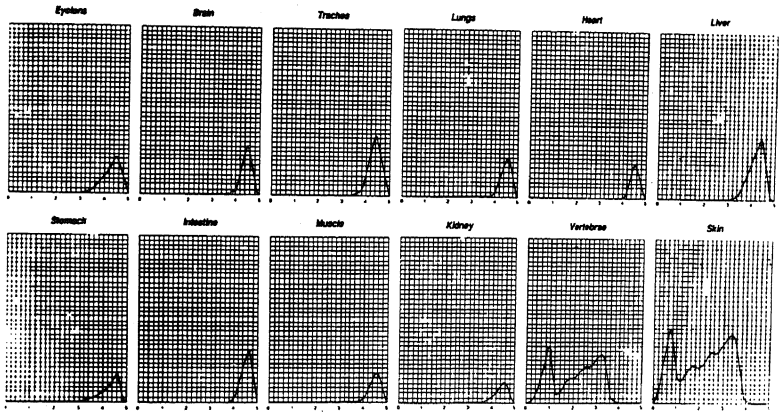


Figure 4. Densitometric scanning of LDH isoenzymes of *Eryx conicus*.

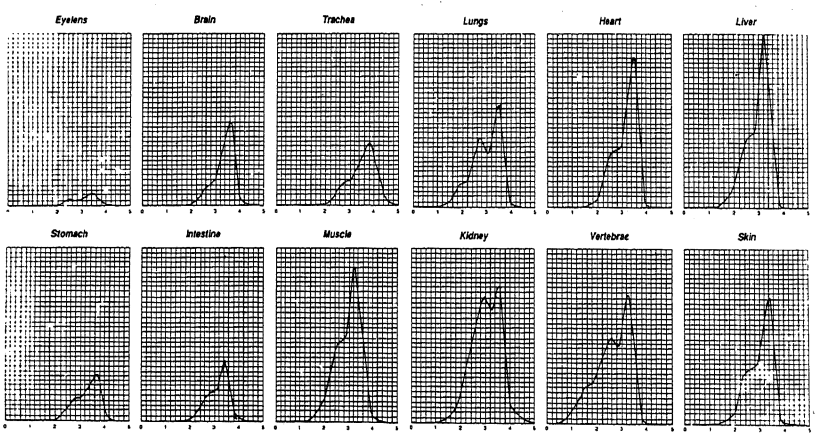


Figure 5. Densitometric scanning of LDH isoenzymes of *Eryx johnii*.

In several vertebrate species, regulatory genes that control the tissue expressions of the LDH-A and LDH-B genes have been detected[35–37]. These observations provide a basis for proposing that the species difference in the tissue patterns of LDH gene expression result from mutational changes at such regulatory genes.

Vertebrate LDH isoenzymes show one- to five-isoenzyme pattern and particularly the two-gene–five-isoenzyme model found in most of the mammals is the full array of LDH-A and LDH-B combination.

All tissues studied contained detectable amount of at least one of the five isoenzymes. Since there seems to be a homology between tissue distribution of isoenzymes in snakes and in other vertebrates, the same subunit designations A and B are used as applied to other vertebrates.

The predominance of LDH-1 (B_4) isoenzyme in all the tissues except vertebrae and skin in *Eryx conicus*, which was not observed in *Eryx johnii*, provides us a clue for identification of *Eryx conicus*.

The predominance of LDH-3 (A_2B_2) and LDH-2 (A_1B_3) in all the tissues except for lungs and vertebrae in *Eryx johnii*, which were not observed in *Eryx conicus*, gives a clear demarcation in establishing species identity.

Here with this authentic electrophoretic evidence, the LDH profile can be used as a specific biochemical marker for establishing species identity in the evolutionary scale of organisms at different levels of biological organization.

Lactate dehydrogenase of rat tissues was selectively inactivated by sodium deoxycholate[38]. Yet another highlight of this investigation is the discovery of sodium deoxycholate as specific inactivator of LDH-5 (A_4) isoenzyme in snake tissues.

1. Markert, C. L. and Moller, F., *Proc. Natl. Acad. Sci.*, 1959, **45**, 753–763.
2. Fisher, S. E., Shaklee, J. B., Ferris, S. D. and Whitt, G. S., *Genetica*, 1980, **52/53**, 73–85.
3. Callegarini, C. and Basaglia, F., *Boll. Zool.*, 1978, **45**, 35–40.
4. Whitt. G. S., *Isoenzymes: Current Topics in Biological and Medical Research* (eds Rattazzi, M. C., Scandalios, J. G. and Whitt, G. S.), Alan R. Liss, New York, 1987, vol. 15, pp. 1–26.
5. Basaglia, F. and Callegarini, C., *Comp. Biochem. Physiol.*, 1988, **B89**, 731–736.
6. Appella, E. and Markert, C. L., *Biochem. Biophys. Res. Commun.*, 1961, **6**, 171–176.
7. Market, C. L., *Science*, 1963, **140**, 1329–1330.
8. Markert, C. L., Sixth Int. Congr. Biochem. Abstracts, 1964, **IV-104**, 320.
9. Epstein, C. J., Carter, M. M. and Goldberger, R. F., *Biochem. Biophys. Acta*, 1964, **92**, 391–394.
10. Markert, C. L., *Ann. NY Acad. Sci.*, 1968, **151**, 14–40.
11. Markert, C. L., *J. Cell Physiol.*, 1968, **72**, 213–230.
12. Markert, C. L. and Whitt, G. S., *Experientia*, 1968, **24**, 977–991.
13. Markert, C. L. and Faulhaber, I., *J. Exp. Zool.*, 1965, **159**, 319–332.
14. Salthe, S. N., Chilson, O. P. and Kaplan, N. O., *Nature*, 1965, **207**, 723–726.
15. Odense, P. H., Allen, T. M. and Leung, T. C., *Can. J. Biochem.*, 1966, **44**, 1319–1326.
16. Markert, C. L., Seventh Int. Congr. Biochem. Abstracts, 1967, p. 828.
17. Whitt, G. S., *Arch. Biochem. Biophys.*, 1970, **138**, 352–354.
18. Dietz, A. A., Lubrano, T. and Rubinstein, H. M., in *Standard Methods in Clinical Chemistry*, Academic Press, New York, 1972, vol. 7, p. 49.
19. Schwantes, M. L. B., *J. Exp. Zool.*, 1973, **185**, 311–316.
20. Markert, C. L., in *Hereditary Developmental and Immunologic Aspects of Kidney Disease* (ed. Metcoff, J.), Northwestern University Press, 1962.
21. Markert, C. L., in *The Harvey Lectures*, Series 59, Academic Press, New York, 1965, pp. 187–218.
22. Markert, C. L. and Ursprung, H., *Dev. Biol.*, 1962, **5**, 363–381.
23. Kettler, M . K., Ghent, A. W. and Whitt, G. S., *Mol. Biol. Evol.*, 1986, **3**, 485–498.
24. Markert, C. L., Shaklee, J. B. and Whitt, G. S., *Science*, 1975, **189**, 102–114.
25. Ferris, S. D. and Whitt, G. S., *Nature*, 1977, **265**, 258–260.
26. Zuckerandl, E., *J. Mol. Evol.*, 1978, **12**, 57–89.
27. Whitt, G. S., in *Proceedings of the Second International Congress of Systematic and Evolutionary Biology* (eds Scudder, G. G. E. and Revel, J. L.), Hunt Institute for Botanical Documentation, Pittsburgh, 1981, p. 271.
28. Whitt, G. S., *Isoenzymes: Current Topics in Biological and Medical Research* (eds Rattazzi, M. C., Scandalios, J. G. and Whitt, G. S.), Allan R. Liss, New York, 1983, vol. 10, p. 1.
29. Whitt, G. S., *Cell Biochem. Funct.*, 1984, **2**, 134–139.
30. Don, M. and Masters, C. J., *Int. J. Biochem.*, 1976, **7**, 215–220.
31. Nadal-Ginard, B., *Proc. Natl. Acad Sci.*, 1976, **73**, 3618–3622.
32. Nadal-Ginard, B., *J. Biol. Chem.*, 1978, **253**, 170–177.
33. Fritz, P. J. and Pruitt, K. M., in *Isozymes: Current topics in Biological and Medical Research* (eds Rattazzi, M. C., Scandalios, J. G. and Whitt, G. S.), Alan R. Liss, New York, 1977, vol. 1, p. 125.
34. Masters, C. J., *J. Biochem.*, 1982, **14**, 685–689.
35. Shows, T. B. and Ruddle, F. H., *Proc. Natl. Acad. Sci.*, 1968, **61**, 574–581.
36. Khlebodarova, T. M. and Serov, O. L., *Biochem. Genet.*, 1980, **18**, 1027–1039.
37. Stock, V. and Gill, T. J., *Biochem. Genet.*, 1983, **21**, 933–941.
38. Lehnert, T. and Berlet, H. H., *Biochem. J.*, 1979, **177**, 813–818.

Received 26 September 1994; revised accepted 5 June 1995

Effect of addition of SO_2 on solid-state fermentation of apple pomace

V. K. Joshi, D. K. Sandhu* and S. Jaiswal

Department of Postharvest Technology, Dr Y. S. Parmar University of Horticulture and Forestry, Nauni, Solan 173 230, India
*Department of Microbiology, Guru Nanak Dev University, Amritsar 143 005, India

Addition of up to 150 ppm of sulphur dioxide (SO_2) in solid-state fermentation (SSF) of apple pomace by *Candida utilis* and *Torula utilis*, and 200 ppm in case of *Saccharomyces cerevisiae* generally enhanced the ethanol concentration, fermentation efficiency and decreased the reducing sugars. But 250 ppm and above of SO_2 proved inhibitory to all the yeasts tried. The titrable acidity remained similar except for a small increase in the treatment with 300 ppm of SO_2. The yeast count decreased after 200 ppm of SO_2. Bacterial and fungal contamination was the highest at 50 ppm and decreased after this level. The yeast counts at 150–200 ppm of SO_2 after 96 h of SSF of apple pomace were quite comparable with those obtained in the steam-sterilized treatment prior to inoculation. Addition of 150–200 ppm of SO_2 can be used as a practical alternative to sterilization in solid-state fermentation of apple pomace.

APPLE pomace is a waste from apple juice production and its disposal into the environment is a serious problem[1,2]. Its utilization by solid-state fermentation for ethanol[3,4] production or concomitant production of ethanol and animal feed has been reported[5]. To have effective solid-state fermentation, a more or less microbe-free environment is a prerequisite. There are always possibilities of contamination from atmosphere and container[6]. Compared to steam sterilization, use of chemicals for this purpose will be far cheaper. Preservatives such as potassium metabisulphite, sorbic acid and dimethyl-dicarbonate (DMDC) can be used as an alternative to pasteurization, filtration and sterilization[7]. In preservation of fruits and vegetable products KMS (providing SO_2) is used extensively[8] and its use in the solid-state fermentation would be cheaper than steam sterilization. Lack of any information on the use of SO_2 in the solid-state fermentation of apple pomace prompted us to carry out this study.

Three yeasts, viz. *Saccharomyces cerevisiae*, *Candida utilis* and *Torula utilis*, were grown in the yeast malt extract broth and used at the rate of 5% to initiate the solid-state fermentation. Apple pomace used was collected from 'hpmc' fruit-processing plant, Parwanoo (HP). Fresh apple pomace was used within 2 h of collection. Potassium metabisulphite (KMS) was added to the apple pomace. Six concentrations of SO_2 ranging from 50 to 300 mg/l (ppm) were tried. Various combinations of apple pomace after addition of SO_2 were not sterilized. A set without SO_2 addition served as a control. Fermentations were carried out in conical flasks. Before inoculation of yeasts in all the flasks and addition of KMS and ammonium sulphate, 1 g of sample from each flask was taken out aseptically for total microbial count, which was termed as 0 h reading. Fermentation was carried out for 4 days. In all the fermentations, the pH and moisture content were kept at the original level[5]. The fermented apple pomace was analysed for ethanol content by the method described by Caputi *et al.*[9]. The total microbial count was determined by the routine pour plate method[10]. Reducing sugars were measured by the DNS method[11], while total sugars were estimated after hydrolysis with acid prior to estimation. Titrable acidity was estimated as per the routine method[12].

Results in Table 1 reveal that in the case of fermentation of apple pomace by *Saccharomyces cerevisiae*, SO_2 concentration up to 200 ppm increased the ethanol production, after which it declined. However, in the case of fermentation by *Candida* and *Torula* ethanol productivity increased up to 150 ppm SO_2 only. In all the cases, 300 ppm of SO_2 proved to be inhibitory as shown by lowest ethanol production. Increasing SO_2 concentration decreased the sugar content in the medium. But the decrease was not exactly proportional to the increase in ethanol produced. Addition of SO_2 might have provided effective aseptic/anaerobic environment,

Table 1. Effect of SO_2 on some physicochemical characteristics of apple pomace fermented by solid-state fermentation

SO_2* (ppm)	Ethanol (% v/v)			Fermentation efficiency (%)			Reducing sugar (%)			Titrable acidity (% MA)		
	Sac	Ca	Tu	Sac	Ca	Tu	Sac	Ca	Tu	Sac	Ca	Tu
0	3.92	3.71	3.93	60	57	60	2.50	2.76	2.45	0.32	0.35	0.32
50	4.11	3.80	3.93	63	58	60	2.24	2.70	2.39	0.35	0.38	0.35
100	4.38	3.99	3.86	67	61	59	2.42	2.62	2.54	0.38	0.42	0.35
150	4.43	4.59	4.10	67	61	59	2.28	2.10	2.40	0.32	0.37	0.26
200	3.98	4.50	3.92	68	49	60	2.26	2.84	2.54	0.32	0.36	0.32
250	3.47	3.21	4.02	53	69	62	2.46	2.34	2.38	0.38	0.38	0.38
300	2.69	3.62	2.42	41	56	37	3.28	3.12	3.24	0.45	0.54	0.54

Sac, *Saccharomyces cerevisiae*; Ca, *Candida utilis*; Tu, *Torula utilis*; MA, Malic acid.
*As potassium metabisulphite.

Table 2. Effect of different concentrations of SO_2 on microbial population (CFU/g) after 96 h of solid-state fermentation

	Inoculated with								
	Saccharomyces cerevisiae			*Torula utilis*			*Candida utilis*		
Treatment	Bacteria	Fungi	Yeast	Bacteria	Fungi	Yeast	Bacteria	Fungi	Yeast
Sterilized	—	—	4.5×10^5	—	—	4×10^5	—	—	4×10^5
Control (no SO_2)	4.20×10^5	4×10^4	8×10^4	3.8×10^5	1×10^4	1.1×10^5	2×10^5	2×10^4	4×10^4
50 ppm	2.80×10^5	—	1×10^5	2.5×10^5	2×10^4	2.2×10^5	1.8×10^5	3×10^4	2×10^5
100 ppm	1.80×10^5	2×10^4	1.2×10^5	1.6×10^5	2×10^4	2.4×10^5	1.5×10^5	5×10^4	2.3×10^5
150 ppm	6×10^4	1×10^4	1.4×10^5	7×10^4	2×10^4	4×10^5	1×10^5	—	3×10^5
200 ppm	5×10^4	—	4.8×10^5	6×10^4	—	4.5×10^5	8×10^4	—	3.7×10^5
250 ppm	4×10^4	1×10^4	2.5×10^5	4.5×10^4	—	1.2×10^5	8×10^4	—	2.1×10^5
300 ppm	2×10^4	—	2.1×10^5	3×10^4	—	5×10^4	6×10^4	—	1.3×10^5

0 h Reading (microbial population) = bacteria 9×10^6 and yeast 1×10^5.

increasing the ethanol content, or could have maintained the ethanol level already produced in the medium by preventing its oxidation or its consumption by yeast cell after completion of fermentation, as has also been observed by Kargi *et al.*[13] using reducing agents such as cysteine and sodium thioglycollate in the ethanolic fermentation. Further, as the product of anaerobic mode of growth increased, concentrations of SO_2 higher than 150 ppm might have decreased the growth, and hence the metabolite, i.e. ethanol. The decrease of ethanol content at 300 ppm is clearly the result of toxicity of SO_2 to the yeasts. SO_2 after 200 ppm suppressing the fermentation is also supported by the findings of Terrell *et al.*[7], wherein 0.4 mM of SO_2 used in various combinations with other preservatives suppressed the fermentation. Fermentation of sugar by the yeast results in ethanol production and, to provide optimum conditions, naturally increases the efficiency of the process. It corroborates with the behaviour of the yeasts used in solid state fermentation of apple pomace.

The titrable acidity of the fermenting medium remained almost similar except for a slight increase in the pomace when 300 ppm of SO_2 was added. It indirectly shows that ethanolic fermentation took place satisfactorily in all the combinations except where 300 ppm of SO_2 was added. From the results it is also clear that trend in the acidities was almost similar in all the yeasts.

The results (Table 2) on the effect of different concentrations of SO_2 on microbial growth revealed that there is similarity in the effect of SO_2 on the growth of yeast, bacteria and fungi. It was found that 250 ppm of SO_2 was toxic to the yeasts (*Saccharomyces*, *Torula*, *Candida*) as there is low yeast count after 200 ppm of SO_2. A large number of bacteria and fungi were found in the case of yeast-inoculated sets of fermentation, indicating that complete sterility, as with steam sterilization, is not achieved with SO_2. With the increase in SO_2 concentration the bacterial and fungal count decreased, and the yeast count of the inoculated yeast increased. It is desirable from the ethanolic fermentation point of view. It has been reported[3] that natural fermentation of apple pomace also produces ethanol and addition of SO_2 might have prevented the activity of natural microflora and allowed the activity of yeast for better ethanol yield by solid-state fermentation. Microbial counts showed that the bacterial population was found maximum in the case of 50 ppm SO_2 in all the three inoculations of yeasts, and after 50 ppm there is considerable decrease in bacterial growth. Fungal growth was less than bacterial growth in all the concentrations of SO_2 tried. These findings suggested that the use of SO_2 can be made to achieve satisfactory solid-state fermentation of apple pomace.

Considering the ethanol yield, sugar content, titrable acidity, effect on different yeasts and their fermentation efficiency and microbial population, it is apparent that addition of 150–200 ppm of SO_2 is advantageous. Concentration of SO_2 higher than this may not be preferred due to the nonuniform effect on one or the other yeast in ethanolic fermentation of apple pomace by solid-state fermentation.

1. Downing, D. L., *Processed Apple Products*, Van Nostrand Reinhold, New York, 1989, p. 433.
2. Joshi, C. and Joshi, V. K., *Indian Food Packer*, 1990, **44**, 56–67.
3. Hang, Y. D., Lee, C. Y. and Woodams, E. F., *J. Food Sci.*, 1982, **47**, 1851–1852.
4. Hang, Y. D. and Woodams, E. F., *Biotechnol. Lett.*, 1984, **6**, 763–764.
5. Joshi, V. K. and Sandhu, D. K., in *Solid State Fermentation* (ed. Pandey, Ashok) Wiley Eastern, New Delhi, 1994, pp. 93–98.
6. Splittstoesser, D. F., *J. Food Prot.*, 1982, **45**, 874–879.
7. Terrell, F. R., Monis, J. R., Johnson, M. G., Gbur, E. E. and Makus, D. J., *J. Food Sci.*, 1993, **58**, 1132–1134.
8. Potter, N. N., *Food Science*, CBS, Delhi, 1987.
9. Caputi, A., Ueda, M. and Brown, J., *Am. J. Enol. Vitic.*, 1968, **19**, 160–165.
10. Harrigan, H. F. and McCance, E. M., *Laboratory Methods in Microbiology*, Academic Press, London, 1966.
11. Miller, G. L., *Anal. Chem.*, 1959, **31**, 426–428.
12. Ranganna, S., *Handbook of Analysis and Quality Control for Fruit and Vegetable Products*, Tata McGraw Hill, New Delhi, 1986.
13. Kargi, F., Curme, J. A. and Sheehan, J. J., *Biotechnol. Bioeng.*, 1985, **27**, 34–40.

Received 28 March 1995; accepted 20 June 1995

Analysis and significance of female reciprocal call in frogs

Debjani Roy, Bijoylakshmi Borah and Amarendra Sarma

Institute of Self-Organising Systems and Biophysics, North Eastern Hill University, Permanent Campus, Shillong 793 022, India

Female vocalization in frogs, feeble as it is, has often escaped the attention of researchers. Most of the work has focused on male vocalization and female phonotaxis experiments. The present study has for the first time compared in detail the temporal and spectral properties of the male advertisement and conspecific female reciprocal calls and focused on the role of female vocalization in the breeding biology of amphibians.

While studying anuran acoustic communication in the frogs of northeast India, female reciprocal calls given in response to the male advertisement call were recorded from *Rana erythraea*, *Rana limnocharis* and *Rana cyanophlyctis*. Once the female responds to the advertising males, more activity is observed in the breeding colony, involving mostly jumping around and across the responding female. The advertisement calls made by the male after the reciprocal call of the female have a higher intensity.

Comparative Fourier analysis of the female reciprocal call and the conspecific male advertisement call showed that the frequency domain of the male call is almost double that of the female call and accordingly there is a shift in the dominant frequency, whereas the spectral pattern consisting of a single or bimodal frequency distribution is common to both.

ANIMALS often produce vocal signals for communication[1–4]. In amphibians vocal signals have proved invaluable in studying systematic relationships[5–7] and the process of sexual selection[8–10]. Males of many amphibian anuran species produce species-specific complex sounds composed of numerous, closely spaced, often harmonically related components with predominant amplitudes that fall into distinct frequency bands, the locations of which are characteristic for each species. Male calls help to advertise their position, to defend their territories and to attract conspecific females to the breeding site, thus acting as an important determinant of mate choice[11–13]. Females recognize the calls of conspecific males, locate them within a chorus of the calling conspecific and heterospecific males and find a mate. Playback experiments have verified the importance of both the gross and fine temporal characters, as well as the spectral domain and spectral pattern of the call in a noisy acoustic environment[11]. Thus, the use of species-specific temporal and spectral character is one strategy for reducing heterospecific interference and maintenance of reproductive isolation[14–16].

Acoustic communication plays a significant role in the reproductive behaviour and breeding biology of amphibians[17]. Anuran vocalization has been classified according to the behavioural context in which the calls are elicited[18–20]. Most of the literature pertaining to anuran acoustic communication deals specifically with species-specific temporal and spectral characters of the male calls[11,14,21–23]. Although female vocalization had been reported as early as 1929, these reports were restricted only to the alarm, release and aggressive calls in *Rana silvatica*[24], *R. pipiens*[25], *R. esculenta*[26], *Colostethus inguanalis*[27] and *Eleutherodactylus coqui*[28].

Reciprocal call given in response to the male advertisement call was reported in the genus *Tomodactylus* based only on behavioural observations, with, no call recordings[29]. Heinzmann[30] working on *Alytes* sp., recorded two types of male calls – mating call and distress call. This work gives a brief description of the oscillogram for the reciprocal calls of the females, with no details of temporal and spectral characteristics. Female calling has also been reported in *R. ridibunda*[31] and *R. virgatipes*[32]. It is noteworthy to mention here that, though sketchy reports on female reciprocal calls existed[29] since 1957, in-depth study to look into the role of female reciprocal call in the breeding biology of amphibians has started only recently. A reason for this could be that the sensation of hearing for humans is most acute in the 2000–5000 Hz frequency range and the female calls occur much below this. Indeed, it is difficult to distinguish and record a female call, which is usually masked by the loud advertisement calls of the conspecific and heterospecific males in the vicinity.

While studying acoustic communication and related behavioural displays in the courting frogs of northeast India, we observed that females of *R. erythraea*, *R. limnocharis* and *R. cyanophlyctis* emitted reciprocal calls in response to male advertisement calls. Due to their feebleness and short call duration, these reciprocal calls would not have been noticed had they not been given in rapid succession, producing at times chirp-like calls, which attracted our attention. We then recorded and analysed the female reciprocal calls.

The calls were recorded during the active breeding period (May to August 1993 and 1994). In 1993 only behavioural observations were made and in 1994 the calls along with behavioural observations were recorded and analysed. The calls were recorded with a unidirectional AKG C451EB shotgun condenser microphone held approximately 40–60 cm away from the calling female. The calls were recorded with a professional SONY WM-D6C cassette recorder and stored on Maxell XLII cassette tapes. Sound pressure level was measured by playing back the isolated female calls on a Philips double cassette player DR920 with playback volume level fixed at 4 and the CYGNET 2021 sound pressure

level meter held approximately 60 cm away from the sound source.

Recorded acoustic stimuli were digitized via a Microsoft analogue-to-digital interface board on to an IBM PC and stored on diskettes. Oscillograms, sonagrams and mean spectra were prepared with a computerized Fourier analysis system (FFT) after passing through bandpass filters and printed with a laser printer.

R. erythraea are mostly arboreal in nature and hence their calls and reproductive behaviour were easier to observe and record in comparison to the calls of *R. limnocharis* and *R. cyanophlyctis*. The latter two species called mostly sitting on land on the edge of a swamp and while hovering on water surface, respectively. Besides, these two species were more easily disturbed, and swam deep into the water at the slightest disturbance. Calls of *R. erythraea* were recorded at Guwahati, Assam (200 m ASL; 91.5'E; 26.1'N). *R. limnocharis* and *R. cyanophlyctis* calls were recorded from colonies being maintained under near-natural condition in the froggery adjacent to the Institute at Shillong, Meghalaya (1515 m ASL; 90.7'E; 24.0'N).

The females of *R. erythraea* (Figure 1 *A*) are usually seen emerging and approaching the calling males after about 30 min of vocal advertisement by the male. When the distance between the calling male and the approaching female reduces to about 10 cm, the male which had been calling till then stops advertising. The female makes feeble reciprocal calls in response to the male advertisement call. This feeble call of the female at times continues up to 1 h, with several bouts of call. The interval between each bout varies from 2 to 5 min, short intervals at the beginning and long intervals towards the end of the call. Only after the female responds to the male advertisement call, the male which had stopped calling on the appearance of the female restarts calling with a higher intensity. Concomitantly, all the other neighbouring calling males also start calling at a higher intensity. More activity is seen amongst the calling males – mostly jumping from one place to another, but all jumps are confined around and across that female, which responded by giving the reciprocal call.

R. limnocharis and *R. cyanophlyctis* (Figure 1 *B, C*) females called while submerged in water, making the feeble reciprocal calls very difficult to hear and record and the behavioural displays difficult to observe. In both the species the females responded to the male advertisement calls. Once the reciprocal calls were emitted, the male calls were louder and more activity was seen amongst the males which were nearer to the responding female.

The reciprocal call of *R. erythraea* is low-pitched, lasting approximately 32 ms. The call sometimes has 3 to 4 subunits, each subunit not lasting more than 10 ms. The call is composed of about 22 pulses. The dominant frequency is at about 1048 Hz, with no harmonics, and the frequency domain extends from about 746 to 1229 Hz, the SPL being 77 dB (Table 1; Figure 1 *a*).

In *R. limnocharis* the calls lasted approximately 61 ms, with 45 pulses. The call had a single dominant frequency at about 1531 Hz, lacking harmonics. The frequency domain is comparatively small, extending from about 1300 to 1767 Hz, the SPL being 70 dB (Table 1; Figure 1 *b*).

In *R. cyanophlyctis* the call duration was approximately 20 ms, composed of 20 pulses, having constant interpulse interval. The frequency distribution was bimodal, with dominant frequency at about 740 Hz and its second harmonic at about 1433 Hz. The frequency domain extends from about 496 to 1746 Hz, the SPL being 65 dB (Table 1; Figure 1 *c*).

Comparative spectral analysis of the female reciprocal calls with male advertisement calls of *R. erythraea* (Roy, unpublished), *R. limnocharis* and *R. cyanophlyctis*[7] shows that female calls are much shorter in duration, with comparatively more pulses, resulting in reduced interpulse interval and pulse overlap. There is a shift in the frequency domain towards the lower range (almost half) in all the three female reciprocal calls, resulting in a shift of dominant frequency as well. Interestingly, in spite of the spectral frequency shift to the lower range, the spectral pattern along with the envelope curve remains unchanged. Like their male counterparts, there is only a single dominant frequency in *R. erythraea* and *R. limnocharis* and bimodal frequency distribution in *R. cyanophlyctis*.

Female vocalization due to its feebleness and the secretive nature of the females has often escaped the attention of researchers. Most works done till now have been on male vocalization and female phonotaxis experiments. This study has for the first time focused on the role of female vocalization in the breeding biology of the amphibians and demonstrated that the female reciprocal calls seem to act as a 'catalyst' for the enhancement of the reproductive activity of the breeding colony.

The feebleness of the call may be due to various reasons: (a) The bigger size of the females compared to the males[5,33]. (b) The absence of vocal sacs in the females; the expandable pouch-like vocal sac present only in the males acts as a sound resonator and radiator[1,19,34]. However, little evidence exists to support this hypothesis and it is now being contemplated that the vocal sacs in the males may have a 'variety of mutually nonexclusive functions'; this is still being worked upon[35,36]. (c) The differences in the structural morphology and tension of the musculature of the vocal cords and larynx and also the tension of the trunk muscles[37–40]. A comparative study on the morphology

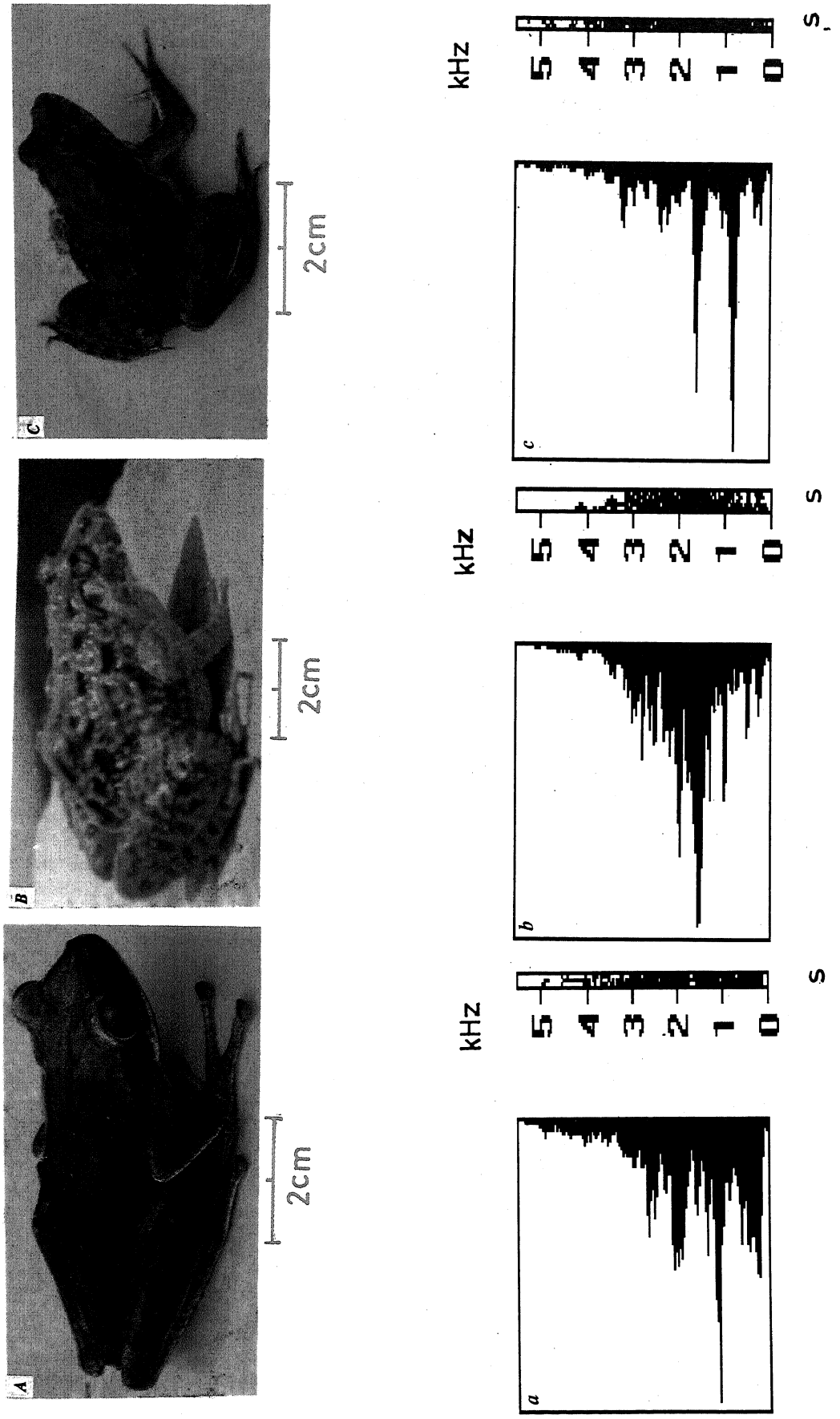
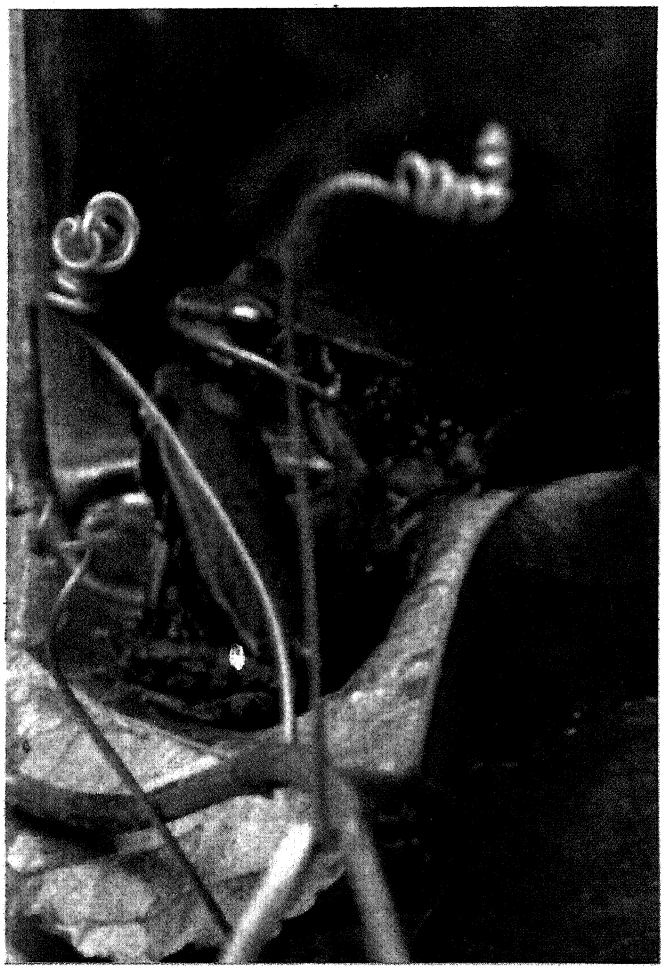


Figure 1. *A*, Female *R. erythraea*; *a*, mean spectrum of female reciprocal call of *R. erythraea*; on the right: the corresponding sonagram (FFT length: 256; overlap: 50%, window: Hamming). *B*, Female *R. limnocharis*; *b*, mean spectrum of female reciprocal call of *R. limnocharis*; on the right: the corresponding sonagram (FFT length: 256; overlap: 50%; window: Hamming). *C*, Female *R. cyanophlyctis*; *c*, mean spectrum of female reciprocal call of *R. cyanophlyctis*; on the right: the corresponding sonagram (FFT length: 256; overlap: 50%, window: Hamming).



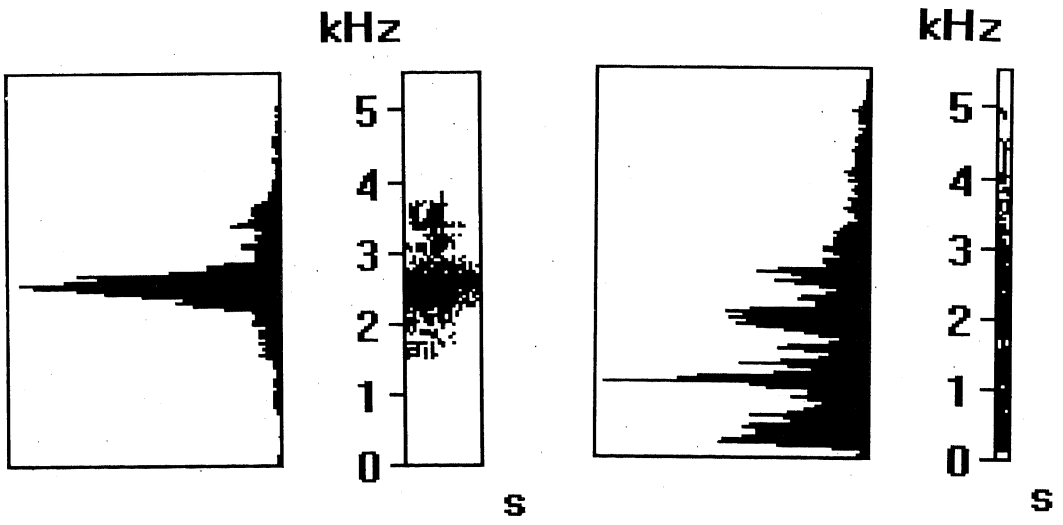


Figure 2. *Top*, Male (left) and female (right) *R. erythraea*. ***Bottom***, Mean spectrum and sonagram of the male advertisement call of *R. erythraea* (left) and mean spectrum and sonagram of the female reciprocal call of *R. erythraea* (right) (FFT length: 256; overlap: 50%, window: Hamming).

Table 1. Comparative analysis of female reciprocal calls and male advertisement calls of *R. erythraea*, *R. limnocharis* and *R. cyanophlyctis*

Species	No. of calls per species	Call duration (ms)	Pulse number	Frequency (Hz)		
				Dominant frequency	Lower limit	Upper limit
R. erythraea						
Female	14	32 ± 9	22 ± 9	1048 ± 105	746 ± 126	1229 ± 114
Male	15	224 ± 4	25 ± 1	2459 ± 42	1495 ± 23	3724 ± 31
R. limnocharis						
Female	14	61 ± 27	45 ± 7	1531 ± 204	1300 ± 224	1767 ± 213
Male	40	503 ± 101	80 ± 13	2144 ± 1254	1576 ± 87	2685 ± 115
R. cyanophlyctis						
Female	12	20 ± 4	20 ± 6	740 ± 36 1433 ± 84	496 ± 57	1746 ± 4
Male	34	615 ± 155	7 ± 1	1651 ± 43 3192 ± 213	1163 ± 124	3907 ± 42

Spectral data on the male advertisement call of *R. erythraea* (Roy, unpublished); spectral data on *R. limnocharis* and *R. cyanophlyctis* from ref. 7.
Values are mean ± SD.

as well as the kinematics of the vocal sacs of the conspecific male and female vocal apparatus is necessary to come to a definite conclusion in this matter.

Bioacoustic analysis of amphibian vocalization has emphasized the species-specific temporal and spectral characteristics of the calls. Shift in the temporal pattern occurs which helps avoid acoustic interference and increase the attractiveness of the call is well known[41,42], but shift in the spectral frequency domain has not been reported. The present study in the three ranid species shows that the frequency domain of the male call is almost double that of the female call and accordingly there is a shift in the dominant frequency, whereas the species-specific spectral pattern of having a single or bimodal frequency distribution remains unaltered (Figure 2). The result of this comparative analysis raises an important question – should the spectral pattern, due to its unaltered species-specific nature, be considered more important than the spectral domain for species identification on the basis of bioacoustic analysis of frog calls?

1. Duellman, W. E. and Trueb, L., in *Biology of Amphibians*, McGraw-Hill, New York, 1986, pp. 87–107.
2. Ewing, A. W., in *Arthropod Bioacoustics*, Cornell University Press, Ithaca, NY, 1989.
3. Farabaugh, S. M., in *Acoustic Communication in Birds* (eds Kroodsma, D. E., Miller, E. H. and Oullet, H.), Academic Press, New York, 1982, vol. 2, pp. 85–124.
4. Otte, D., in *How Animals Communicate* (ed. Sebeok, T. A.), Indiana University Press, Bloomington, 1977, pp. 334–361.
5. Blair, A. P., *Q. Rev. Biol.*, 1964, **39**, 334–344.
6. Littlejohn, M. J., in *Evolution and Speciation* (eds Atchley, W. R. and Woodruff, D. S.), Cambridge University Press, Cambridge, 1981, pp. 298–334.
7. Littlejohn, M. J. and Watson, G. F., *Annu. Rev. Ecol. Sys.*, 1985, **16**, 85–116.
8. Gerhardt, H. C., *Am. Zool.*, 1982, **22**, 581–595.
9. Ryan, M. J., *Trends Ecol. Evol.*, 1991, **6**, 351–355.
10. Sullivan, B. K., *Gr. Basin Nat.*, 1985, **45**, 688–696.
11. Gerhardt, H. C., in *The Evolution of the Amphibian Auditory System* (eds Fritzsch, B., Ryan, M., Wilczynski, W., Hetherington, T. and Walkowiak, W.), John Wiley, New York, 1988, pp. 455–483.
12. Rand, A. S., in *The Evolution of the Amphibian Auditory System* (eds Fritzsch, B., Ryan, M., Wilczynski, W., Hetherington, T. and Walkowiak, W.), John Wiley, New York, 1988, pp. 415–432.
13. Ryan, M. J., in *The Evolution of the Amphibian Auditory System* (eds Fritzsch, B., Ryan, M., Wilczynski, W., Hetherington, T. and Walkowiak, W.), John Wiley, New York, 1988, pp. 637–678.
14. Drewry, G. E. and Rand, A. S., *Copeia*, 1983, **1984**, 941–953.
15. Duellman, W. E. and Pyles, R. A., *Copeia*, 1983, **1983**, 639–649.
16. Narins, P. M. and Zelick, R. D., in *The Evolution of the Amphibian Auditory System* (eds Fritzsch, B., Ryan, M., Wilczynski, W., Hetherington, T. and Walkowiak, W.), John Wiley, New York, 1988, pp. 511–536.
17. Wells, K. D., in *The Evolution of the Amphibian Auditory System* (eds Fritzsch, B., Ryan, M., Wilczynski, W., Hetherington, T. and Walkowiak, W.), John Wiley, New York, 1988, pp. 433–454.
18. Bogert, C. M., in *Animal Sounds and Communication* (eds Lanyon, W. E. and Tavolga, W. N.), American Inst. Biol. Sci. Publ., 1960, pp. 137–320.
19. Littlejohn, M. J., in *The Reproductive Biology of Amphibians* (eds Taylor, D. H. and Guttman, S. I.), Plenum Press, New York, 1977, pp. 263–294.
20. Wells, K. D., in *The Reproductive Biology of Amphibians* (eds Taylor, D. H. and Guttman, S. I.), Plenum Press, New York, 1977, pp. 233–262.
21. Roy, D. and Elepfandt, A., *J. Biosci.*, 1993, **18**, 381–393.
22. Roy, D., *J. Biosci.*, 1994, **19**, 629–644.
23. Schneider, H., Sinsch, U. and Sofianidou, T. S., *Z. Zool. Syst. Evolutionsforsch.*, 1993, **31**, 47–63.
24. Noble, K. G. and Ferris, E. F., *Le Conte. Am. Mus.*, 1929, **363**, 1–17.
25. Noble, K. G. and Aronson, L. R., *Bull. Am. Mus. Nat. Hist.*, 1942, **80**, 127–142.
26. Wahl, M., *Oecologia*, 1969, **3**, 14–65.
27. Wells, K. D., *Behav. Ecol. Sociobiol*, 1980, **6**, 199–209.
28. Stewart, M. M. and Rand, A. S., *Copeia*, 1991, **1991**, 1013–1024.
29. Dixon, J. R., *Texas J. Sci.*, 1957, **9**, 379–409.

30. Heinzmann, U., *Oecologia*, 1970, **5**, 19–55.
31. Frazer, J. F., in *Reptiles and Amphibians in Britain*, Collins, London, 1983.
32. Given, M. F., *Herpetologica*, 1987, **43**, 467–481.
33. Loftus-Hills, J. J., *Aust. J. Zool.*, 1973, **21**, 353–367.
34. Watkins, W. A., Baylor, E. R. and Bowen, A. T., *Copeia*, 1970, **1970**, 558–561.
35. Dudley, R. and Rand, A. S., *Copeia*, 1991, **1991**, 460–470.
36. Rand, A. S. and Dudley, R., *Physiol. Zool.*, 1993, **66**, 793–806.
37. Taigen, R. L., Wells, K. D. and Marsh, R. L., *Physiol. Zool.*, 1985, **58**, 719–726.
38. Marsh, R. L. and Taigen, T. L., *Am. J. Physiol.*, 1987, **252**, 768–793.
39. Given, M. F. and Mckay, D. M., *Copeia*, 1990, **1990**, 863–867.
40. Given, M. F., *Anim. Behav.*, 1993, **46**, 1139–1149.
41. Rand, A. S., and Ryan, M. J., *Z. Tierpschol.*, 1981, **40**, 279–299.
42. Sullivan, B. K. and Malmos, K. B., *Herpetologica*, 1994, **50**, 146–156.

ACKNOWLEDGEMENTS. We thank Profs R. Gadagkar and V. Nanjundiah for their comments on the manuscript; the Z.S.I., Shillong, for the identification of *Rana erythraea* and the Department of Science and Technology, New Delhi, for financial support.

Received 27 March 1995; revised accepted 12 May 1995

Yttrium-, niobium- and zirconium-rich rhyolite dykes of Dhorio Nes area, district Jamnagar, Gujarat, India

P. B. Maithani, Rahul Banerjee, Navin Goyal, S. Ramachandran and Rajendra Singh*

Atomic Minerals Division, Sharada Building, R. V. Desai Road, Baroda 390 001, India
*Atomic Minerals Division, AMD Complex, Begumpet, Hyderabad 500 016, India

Anomalously high contents of Nb (818 ppm), La (173 ppm), Y (305 ppm) and Zr (> 1000 ppm) are reported for the first time in the rhyolite dykes/veins spatially associated with the Deccan basalts in Alech hills around Dhorio Nes, district Jamnagar, Gujarat. The enrichment of these elements is due to influx of alkalies and volatiles in rhyolites emplaced along the western continuity of E–W-trending Son–Narmada rift.

THE Deccan Traps, covering about one-sixth of India's total land surface, comprise mostly of tholeiitic basalt. Acidic, alkalic and mafic alkalic rocks are found in the western part of the basalt plateau in western India in parts of Kutch, Saurashtra, Lower Narmada Valley and Bombay coast.

The Saurashtra region of Gujarat is known for a number of acid igneous rock complexes along the western extension of the Son–Narmada lineament. These include the Alech–Barda hills, Osham hills, Girnar hills and Chogat–Chamardi hills[1]. These acidic rocks developed as an independent magma suite with rhyolite as a predominant rock type within the complex[2,3], or as an end product of differentiation[4].

In Alech–Barda hills, the acid igneous rocks comprise pitchstone, rhyolite, granophyre and diorite. The rhyolites occur as dykes and veins within the basalt, and also as flows of rhyolitic crystal tuffs. Two prominent pink- to buff-coloured, highly fractured E–W-trending rhyolite dykes were noticed in the south and southwest of Dhorio Nes (69°59′44″ : 21°48′27″, Toposheet no. 41 G/13) in the Alech hill area, showing higher than background radioactivity (Figure 1). They extend over a length of 500 m with width varying from 10 to 30 m. They consist mainly of quartz, sanidine and glass along with some opaques and exhibit vitrophyric texture. The groundmass in the rhyolites is essentially composed of a glassy material along with crystallites of feldspars. The chemical data of these rocks along with the data of Pavagarh and Amreli rhyolites and the general abundance in rhyolites is given in Table 1. The samples show normative corrundum, peraluminous ($A/CNK = 1.52$), miaskitic ($AI = 0.64$) nature and enrichment of potash over soda.

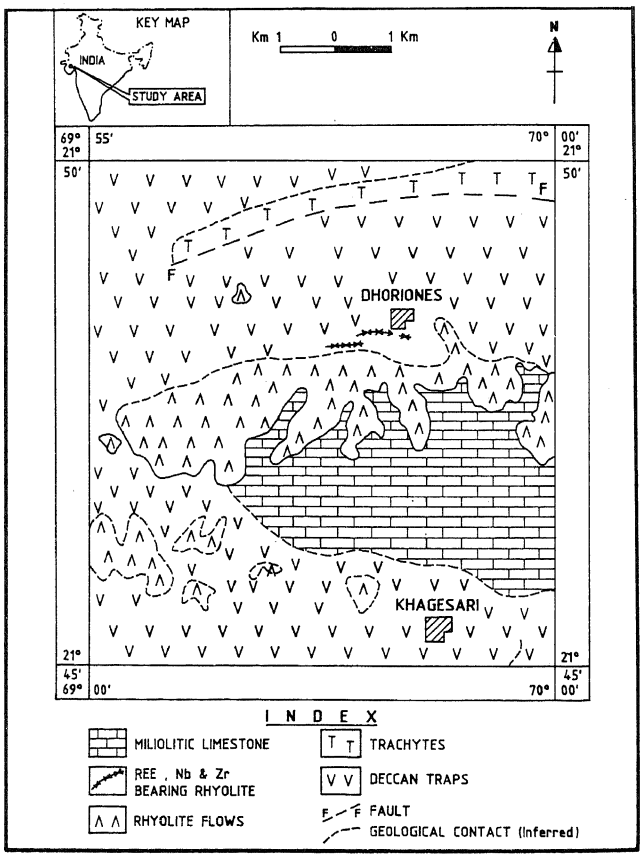


Figure 1. Geological map showing rare-earths-, niobium- and zirconium-rich rhyolites of Dhorio Nes area, district Jamnagar, Gujarat.

Table 1. Major oxide and trace element data of rhyolites of Dhorio Nes, Pavagarh hills and Amreli, Gujarat

	1	2	3	4	*X*	SD	C%	5	6	7
Major oxides										
SiO_2	72.69	71.63	71.63	71.88	71.95	0.50	0.69	72.29		72.77
Al_2O_3	12.53	11.36	12.53	11.81	12.05	0.57	4.73	12.49		13.33
TiO_2	0.18	0.15	0.17	0.21	0.17	0.025	14.70	0.30		0.29
Fe_2O_3	3.14	4.47	4.13	3.20	3.73	0.66	17.69	0.73		1.40
FeO	0.06	0.06	0.04	0.30	0.115	0.12	104.34	2.49		1.02
MgO	0.15	0.07	0.08	0.10	0.10	0.03	30.00	0.14		0.38
CaO	0.13	0.12	0.10	0.13	0.12	0.014	11.66	0.72		1.22
Na_2O	1.24	1.63	1.71	2.34	1.73	0.45	26.01	3.93		3.34
K_2O	6.17	3.83	3.91	4.29	4.55	1.09	23.95	4.83		4.58
MnO	0.04	0.02	0.02	0.02	0.025	0.01	40.00	0.07		0.07
P_2O_5	0.01	< 0.01	0.02	0.01	0.011	0.006	54.54	0.06		0.10
LOI	2.69	2.90	2.53	1.70	—	—	—	—		—
A/CNK	1.40	1.63	1.71	1.35	1.52	—	—	0.96		1.07
AI	0.70	0.59	0.57	0.72	0.64	—	—	—		—
Trace elements										
Y	276	164	331	337	277.00	80.17	28.94	92	52	10–50
Zr	> 1000	> 1000	> 1000	> 1000	—	—	—	660	380	200–400
Sn	41	34	47	38	40.00	5.47	13.67	—	10	< 5
Nb	510	852	1000	565	731.75	233.37	31.89	83	18	5–40
La	65	20	128	251	116.00	100.30	86.46	—	49	20–50
U	40	30	30	50	37.50	9.57	25.52	—	—	4–5
Th	110	130	130	130	125.00	10.00	8.00	—	—	20

Columns 1–4: Major oxide (in wt%) and trace element (in ppm) data of rhyolites of Dhorio Nes, district Jamnagar, Gujarat.
Column 5: Major oxide (in wt%) and trace element (in ppm) data of rhyolites of Pavagarh hills ($n = 1$), district Panchmahals, Gujarat[6].
Column 6: Average of trace element (in ppm) data of rhyolites of Amreli ($n = 12$), Gujarat[7].
Column 7: General abundance of major oxides and trace elements in rhyolites[8,9].

Table 2. Trace element data (range) of basalt, rhyolitic flows and rhyolitic dykes of the Dhorio Nes area (all values in ppm)

Trace element	Basalt		Rhyolitic flows	Rhyolitic dykes
	$n = 4$	General abundance*	($n = 3$)	($n = 14$)
Y	50–115	25	110–168	114–401
Zr	259–409	140	607–967	> 1000
Sn	n.d.–10	1.5	n.a.	30–47
Nb	25–26	19	118–151	510–> 1000
La	< 50	17	< 50–148	42–287
Ti	1434–7760	13800	2343–2562	1054–1434
V	16–89	250	19–27	10–44
Cr	22–66	170	19–30	17–74
Mn	709–1485	1500	1122–1289	71–754
Co	13–28	48	< 5	< 5
Ni	10–41	130	10–13	10–39
Cu	12–168	87	8–9	10–28
Ga	14–19	17	19–22	21–34
Pb	< 10	6	< 10	11–81

n.a. = Not analysed.
n.d. = Not detected.
*Of trace elements in basalt[10].

Unusually high content of Nb (818 ppm), La (173 ppm), Y (305 ppm), Sn (36 ppm), Zr (> 1000 ppm, average 2090 ppm for 2 samples – XRF analysis) and U (35 ppm) were recorded in the rhyolite dykes ($n = 14$) of Alech hills around the Dhorio Nes area of Jamnagar district, Gujarat (Figure 1), which is higher than the general abundance of these elements in acidic rocks. The anomalous content of these elements in the rhyolite dykes of Dhorio Nes is due to the presence of brownish-red- to blood-red-coloured, tetragonal- to square-shaped xenotime and rutile, in addition to the monazite and ilmenorutile, as observed under the microscope. Besides, zircon is also identified by X-ray diffraction studies of heavies separated from rhyolites. The detailed mineralogical studies of slime and light fraction are under progress using electron probe micro analyser (EPMA).

The upper mantle source region may have been enriched or metasomatized to give rise to basalts with anomalous values (Table 2), which in turn fractionated to minor rhyolites with anomalous Y, Nb and Zr values. The Narmada–Son rift, which has been reactivated several times during the geological history, probably acted as a locale for the concentration of these trace elements. The mantle-derived CO_2-rich fluids from this rift also contributed to the trace element pattern observed in the rhyolites. It is pertinent to mention here that there are many carbonatite–alkaline complexes along the Narmada–Son rift[5].

Significant amounts of U, Nb, La, Y and Sn in rhyolites of Dhorio Nes makes Alech hills, a possible potential area for exploration of these metals. Detailed work on the find is underway and will be published shortly.

1. Fedden, F., *Mem. Geol. Surv. India*, 1884, **21**, 1–64.
2. Sukeshwala, R. N., *J. Geol. Soc. India*, 1982, **23**, 13–18.
3. De, Aniruddha, *Mem. Geol. Soc. India*, 1982, **3**, 327–339.
4. Wakhaloo, S. N., Symposium on Upper Mantle Project, Hyderabad, Session VI, 1967.
5. Bose, M. K., Seminar on Geodynamics of Himalayan Region, Hyderabad, 1973, 173–181.
6. Alexander, P. O., *J. Geol. Soc. India*, 1980, **21**, 453–457.
7. Maithani, P. B., Banerjee, R. and Reddy, N. R., Unpub. Annual Report, Atomic Minerals Division, 1993, Hyderabad, p. 43.
8. Wedepohl, K. H., *Handbook of Geochemistry*, Springer, New York, 1969.
9. Daly, R. A., *Igneous Rocks and the Depth of the Earth*, McGraw Hill, New York, 1933, pp. 9–28
10 Turekian, K. K. and Wedepohl, K. H., *Bull. Geol. Soc. Am.*, 1961, **72**, 171–192.

ACKNOWLEDGEMENTS. Analytical support from the Spectrograph and XRF Laboratories, Atomic Minerals Division, Hyderabad, and Chemistry Laboratory, Atomic Minerals Division, Baroda, is acknowledged with thanks. We thank Shri K. K. Dwivedy, Director, Atomic Minerals Division, for permission to publish this paper.

Received 24 November 1994; revised accepted 8 May 1995

Evaluation of efficacy of bricks as geochemical monitor of atmospheric heavy-metal pollution

Rohit Shrivastav, Sanjay Kumar Mathur, Shobhit Shrivastav*, M. M. Shrivastav, Sahab Das and Satya Prakash

Department of Chemistry, Dayalbagh Educational Institute, Dayalbagh, Agra 282 005, India
*Present address: Department of Chemistry, University of Allahabad, Allahabad 211 002, India

Geochemical monitoring of atmospheric heavy-metal pollution is so far limited to the use of peat, ice deposits and aquatic sediments as sample type. These systems are open, dynamic and easily affected by climatic variations. In contrast, bricks, which are more compact, can act as a better geochemical monitor. Heavy-metal (Cu, Zn, Ni, Cr, Pb) analysis in scores of soil and baked/unbaked brick samples collected from a large area in and around Agra City indicates approximately similar flux in soils and bricks and insignificant fractionation during brick construction of these metals. The concentration of analysed metals in the core of bricks remained unaffected even by substantial amounts of acidic/alkaline rain under natural conditions. Thus, the feasibility of a new, hitherto untested, concept of bricks acting as a geochemical monitor of atmospheric heavy-metal pollution has been systematically explored.

ATMOSPHERIC heavy metals interact with terrestrial environment continuously. From the standpoint of mass, their major amount is absorbed by plants and soils[1,2]. Chronological changes in the heavy-metals flux of terrestrial objects, mainly as a result of atmospheric depositions, reflect the impact of human growth on accumulation of toxic metals on the global ecosystem over the years. Studies of the practical applications of geochemical monitoring have been carried out mainly in Northwest Europe, North America and the remote polar regions, particularly in relation to the historical aspects of pollution. Records from peat deposits and lake sediments in Europe document the growth and development of industry there over the last 2000 years[3–8]. Reconstructed historical deposition rates demonstrate the severity of pollution in urban areas of northern Britain in the eighteenth and nineteenth centuries, whereas analysis of recent sediments reveals a variable pattern of present-day pollution, with some areas experiencing increasing levels of deposition and others decreasing ones[9–11]. In North America deposition records reveal a much shorter history of atmospheric pollution and document the rapid spread of heavy metals during the mid and late nineteenth century, both on a local and on a regional scale[12–16]. For most of the other parts of the world, no such detailed information exists and studies are limited mainly due to lack of suitable sample types, where the original heavy-metals compositions remain preserved unperturbed over long periods of time. Peat, ice deposits and aquatic sediments, most extensively used so far in monitoring the atmospheric depositions of heavy metals are open, dynamic systems and can be easily affected by climatic changes.

The use of baked bricks in the construction of buildings is an age-old practice all over the world. The commercial process of construction and baking of bricks has not changed much over the years, and the stuff used in brick manufacture is primarily soil derived from the top 1–1.5 m of the earth, a portion where atmospheric depositions of heavy metals are most likely to be retained[2]. The temperature of bricks during commercial baking rises up to $\approx 700^{\circ}C$, at which most of the soil-bound heavy metals possibly remain intact in soil matrix. Hence, bricks of any specific period of time, if kept preserved from the effects of rain and radiations, can be regarded as representative samples of soil of that period, from the viewpoint of heavy-metals composition. Bricks, which can be effectively used for chronological monitoring of atmospheric heavy-metal depositions in soil and which are likely to be better than so far used sample types, viz. ice, peat and sediments, have eluded the attention of researchers till now. This paper deals with a systematic evaluation of the effectiveness of bricks for this purpose.

From an approximately $10^3\ km^2$ area in and around Agra City the following samples were collected: (i) baked and unbaked (65 each) brick samples from

different brick kilns; (ii) bulk soil samples from 35 places after removing the top 7–10 cm layer of the earth; (iii) bulk soil samples from 5 places (wide apart) after every 30 cm depth (up to 1.5 m). The soil samples were air-dried (at 35°C), mildly ground and mixed well; the texture of few representative samples was determined by the wet sieving procedure. The inner core (≈ 50 g) of each brick sample was taken out by removing the external layers in a mechanical cutter and ground to powder form avoiding external contaminations. The grinding was done till the final texture of the brick powder was almost identical to the texture of the soil. The acid extracts of powdered brick and soil samples, generated under mild, semidrastic and drastic acid treatments (detailed below), were quantitatively analysed for Cu, Zn, Ni, Cr and Pb using the standard methods[17], employing Perkin Elmer 2380 flame atomic absorption spectrometer at appropriate wavelengths.

For mild treatment the sample (6 g) was treated with 10 ml HCl (3 M) and 10 ml HNO_3 (3 M) for 24 h and then boiled till the volume was reduced to 3–5 ml. After adding 20 ml of 0.01 M HCl, the contents were further heated for 15 min, cooled, filtered through acid-washed Whatman 42 filter paper and the extract was collected for analysis.

For moderate treatment the sample (6 g) was treated, in a teflon vessel, with 10 ml HCl (6 M), 10 ml HNO_3 (6 M) and 1 ml HF (48%) for 24 h, followed by slow heating at 45–60°C till the volume was reduced to 3–5 ml. In order to ensure the complete removal of unreacted HF, 20 ml of 0.1 M HCl was added and the contents were further heated at elevated temperature (65–80°C) till the volume was again reduced to 3–5 ml, which was then transferred to a glass vessel using 20 ml of 0.01 M HCl as the washing solution. The contents were boiled till the volume was reduced to half, cooled, filtered as above and the extract was collected.

For drastic treatment the sample (6 g) was treated, in a teflon vessel, with 5 ml HCl (12 M), 5 ml HNO_3 (16 M) and 10 ml HF (48%) for 24 h and then heated gently at 45–60°C till the volume was reduced to 3–5 ml. Adding 5 ml each of HCl (12 M) and HNO_3 (16 M), the contents were further heated at 65–80°C till the volume was again reduced to 3–5 ml. At this stage the complete removal of unreacted HF was ensured. Using 20 ml of water as the washing solution, the contents were transferred to a glass vessel, boiled till the volume was reduced to half, cooled, filtered as above and the extract was collected.

The final volume (after washings) and pH of all the extracts were adjusted, respectively, to 50 ml and 3–4 by suitable additions of HCl (0.01 M), NaOH (0.01 M) and water. Metal concentrations were determined in reagent-blank extracts also and the values were subtracted from the main data. The extract samples were stored in acid-washed polythene bottles before analysis. Double-distilled deionized water and AR grade reagents were used throughout the study.

The textures of 20 randomly chosen samples each of powdered bricks and soils were determined (Table 1). The results indicate approximate similarity of the two systems. The average concentrations of Cu, Zn, Ni, Cr and Pb in soil and baked/unbaked brick samples, as determined by the analysis of extracts generated under mild, moderate and drastic treatments, are presented in Table 2. Table 3 shows the results of some laboratory experiments to observe the effects of rain (of varying pH) on heavy-metal concentrations in the core of baked bricks while the bricks are exposed to natural conditions at 30–35°C.

Soil concentrations of metals are in the range of earlier reported values of some Indian soils[18]. Extracts of soil and brick samples generated through mild and moderate extraction procedures show a wide variation in their metal concentrations, whereas drastic-treatment data indicate approximately similar concentrations of metals in soils and baked/unbaked bricks. It is also clear that except for Cr, which shows a slight increase, the concentrations of the remaining four metals in soils of the studied area are not significantly changing with depth (ranging from 0.1 to 1.5 m). Metal concentrations in the core of baked bricks remain unaffected by substantial amounts of acidic/alkaline rain and radiations. However, prolonged exposure of bricks to these results in some changes.

The close association of soil-bound heavy metals with soil matrix, largely composed of silicates, is well known and might further increase during brick construction, depending upon the maximum temperature reached during baking. Hence, determination of metal concentrations in soil and brick extracts depends on the effectiveness of the extraction technique in breaking/dissolving the soil matrix. The wide range of fluctuations in data (Table 2) under mild and moderate treatments indicates inadequacy of extraction procedures in breaking the matrix, and only the drastic treatment, with 95–100% dissolution of the sample, seems to give satisfactory results.

The temperature of bricks during commercial baking

Table 1. Texture of powdered brick and soil samples

Particle size (mm)	Percentage amount*	
	Soil	Powdered brick
60–2	0.36–0.42	0.41–0.49
2–0.2	36.13–40.50	38.10–43.21
0.2–0.06	27.91–34.12	31.63–39.00
0.06–0.006	14.93–19.23	11.54–16.24
0.006–0.002	3.04–7.33	3.02–7.10
< 0.002	3.52–10.10	1.09–6.25

*Based on measurements in 20 randomly chosen samples.

Table 2. Mean concentration (mg kg^{-1}) of metals in soil and brick samples

Extraction technique		Concentration				
		Cu Mean ± SD	Zn Mean ± SD	Ni Mean ± SD	Cr Mean ± SD	Pb Mean ± SD
Drastic	*(0.07–0.1)	1.9 ± 0.15	4.8 ± 0.39	5.9 ± 0.49	3.6 ± 0.33	5.0 ± 0.45
	*(0.3)	2.0 ± 0.22	5.0 ± 0.40	5.8 ± 0.50	3.8 ± 0.35	4.9 ± 0.50
	*(0.6)	1.9 ± 0.18	4.7 ± 0.35	6.2 ± 0.50	3.9 ± 0.30	5.1 ± 0.45
	*(0.9)	2.2 ± 0.19	4.8 ± 0.35	6.3 ± 0.49	4.2 ± 0.32	5.1 ± 0.50
	*(1.2)	2.1 ± 0.20	4.9 ± 0.37	5.9 ± 0.48	4.6 ± 0.30	5.2 ± 0.40
	*(1.5)	2.0 ± 0.17	4.9 ± 0.37	5.9 ± 0.51	4.6 ± 0.33	5.0 ± 0.44
	†	2.1 ± 0.19	4.8 ± 0.40	6.3 ± 0.50	3.9 ± 0.35	5.2 ± 0.50
	‡	2.0 ± 0.21	4.8 ± 0.38	6.0 ± 0.49	4.0 ± 0.33	4.8 ± 0.46
Moderate	*(0.07–0.1)	1.9 ± 0.81	3.9 ± 1.28	4.2 ± 1.83	2.3 ± 1.52	2.9 ± 1.79
	†	1.4 ± 0.93	3.0 ± 1.35	3.6 ± 1.64	3.3 ± 1.69	3.5 ± 1.92
	‡	1.1 ± 0.80	2.6 ± 1.44	3.7 ± 1.53	2.9 ± 1.67	2.5 ± 1.84
Mild	*(0.07–0.1)	0.8 ± 0.32	0.9 ± 0.71	1.2 ± 0.87	0.9 ± 0.76	1.5 ± 0.89
	†	0.6 ± 0.41	0.7 ± 0.63	0.9 ± 0.80	1.3 ± 0.66	1.1 ± 0.92
	‡	0.4 ± 0.33	1.2 ± 0.94	1.3 ± 0.91	0.8 ± 0.78	0.9 ± 0.81

*Soil; †unbaked bricks; ‡baked bricks.
Values in parenthesis represent depth in m, from top, at which the sample was collected.

Table 3. Mean concentration (mg kg^{-1}) of metals in the core of baked bricks, after their exposure to intermittent simulated rain (100 ml day^{-1}) under natural conditions at 30–35°C (drastic-treatment data)

Rainfall		Concentration*				
Duration in days	pH	Cu Mean ± SD	Zn Mean ± SD	Ni Mean ± SD	Cr Mean ± SD	Pb Mean ± SD
2	7.5	2.2 ± 0.17	4.9 ± 0.40	6.2 ± 0.50	3.9 ± 0.33	5.1 ± 0.43
5	7.5	2.3 ± 0.19	4.7 ± 0.37	6.0 ± 0.51	3.9 ± 0.35	5.1 ± 0.47
10	7.5	2.0 ± 0.18	4.8 ± 0.39	5.9 ± 0.47	3.5 ± 0.31	4.9 ± 0.46
2	6.0	2.0 ± 0.21	4.8 ± 0.35	6.0 ± 0.48	4.0 ± 0.34	5.2 ± 0.50
5	6.0	2.2 ± 0.22	4.6 ± 0.37	6.2 ± 0.51	3.8 ± 0.35	5.1 ± 0.48
10	6.0	1.9 ± 0.22	4.5 ± 0.41	6.1 ± 0.49	3.5 ± 0.33	4.8 ± 0.49
2	5.0	2.1 ± 0.19	4.9 ± 0.41	5.9 ± 0.47	4.1 ± 0.35	5.1 ± 0.50
5	5.0	1.9 ± 0.22	4.6 ± 0.39	5.8 ± 0.50	3.8 ± 0.35	4.8 ± 0.50
10	5.0	1.9 ± 0.21	4.4 ± 0.41	5.6 ± 0.50	3.6 ± 0.34	4.6 ± 0.51

*Average of 40 samples.

rarely exceeds 700°C, and the existing vaporization/dissociation temperature data[19] of different metal salts suggest that the major amount of most of the soil-bound heavy metals possibly remains intact in the soil matrix at this temperature. This contention is supported by drastic-treatment results (Table 2), indicating approximately similar concentrations of metals in soils, unbaked bricks and baked bricks. The negligible fractionation of metals during brick construction and to some extent insignificant effect of rain and radiation on heavy-metal concentrations in the core of bricks (Table 3) make it feasible to consider bricks as representative samples of the parent soil from the viewpoint of heavy-metals composition. The compactness and easy availability are other distinct advantages in using bricks as a geochemical monitor, rather than ice, peat and sediments.

From the viewpoint of plant nutrition, soil layers at 0.1–1m depth from the top are the most important because mineral-absorbing zones of plant roots generally lie in this region. Changes in concentrations and mobility of metals in this region have a direct consequence on the living systems, and the entry of some toxic metals in the food chain cannot be ruled out. Incidentally, in brick construction also, this portion of soil is mainly used. Thus, bricks act not only as representative samples of the parent soil but also as samples of biologically most important layers of the top soil.

Heavy metals enter the soil through many sources, viz. atmospheric depositions, agricultural activities, industrial and local emissions. By analysing bricks of different ages derived from interior and well-preserved portions of old buildings/monuments for different metals and presenting the data on time scale, it is possible to trace the past and present accumulation rates of metals in soils of a particular area. Moreover, with the help of statistical tools the identification of probable sources of these metals and the estimation of their contributions

is possible. Analysis of samples collected from remote areas, where local effects are negligible, will reflect the rates of global inputs through atmospheric depositions. In contrast to approximate and lengthy stratigraphic, biostratigraphic, radiometric and geomagnetic techniques used for dating aquatic sediments, peat and ice deposits, the dating method for bricks is limited to determining the construction time of the building from which the brick sample is derived. Most parts of the world have witnessed much of the industrial and urban growth only during the last 200 years and brick samples of this duration can be easily obtained and accurately dated. Further, bricks being heavy construction material are never transported over long distances due to economic reasons and remain representative samples of soil of the area of their collection.

Though our study has shown both bricks and soil as true representatives of each other in respect of concentrations of all the five metals analysed, this trend, far from being general, is bound to be dependent on soil characteristics of the particular area and minor variations in brick-making process from place to place. In the event of certain metals showing considerable fractionation during brick construction, the ratio x ($x = [\text{metal}]_{\text{soil}}/[\text{metal}]_{\text{baked brick}}$) may be determined by analysing a large number of present-day samples. With some approximations, the average values of x (for different metals) can be utilized in calculating metal concentrations in soils of different ages from the metal concentrations data of old bricks. The present study, thus, leads to the following conclusions:

1. From heavy-metal concentrations viewpoint, except for few volatile metals, bricks derived from interior and well-preserved parts of buildings can be regarded as representative samples of the parent soil.

2. To some extent, rains and radiations do not change metal concentrations in the core of bricks.

3. Due to greater compactness and easy availability of samples, bricks are likely to be a better geochemical monitor than ice, peat and sediments and can be used in geochemical monitoring of heavy-metal pollution. However, this technique has to be used judiciously for monitoring atmospheric heavy-metal pollution of soil in historical times due to anthropogenic activities in the past, because the parent soil of a brick collected from a specific area may have inherently high heavy-metal content due to natural causes also.

Finally, it should be stated that the conclusions derived here are based on the study restricted to the city of Agra and its suburbs only. More such studies, in other parts of the world, are needed to understand the exact potential of bricks as a geochemical monitor.

1. Hill, A. C., *J. Air Poll. Contr. Assoc.*, 1971, **21**, 341–346.
2. Bohn, H. L., *J. Environ. Quality*, 1972, **1**, 372–377.
3. Aston, S. R., Bruty, D., Chester, R. and Padgham, R. C., *Nature*, 1973, **241**, 450–451.
4. Hamilton, T. J., *Environ. Sci. Technol.*, 1979, **13**, 693–697.
5. Hamilton, T. J., *Environ. Technol. Lett.*, 1983, **4**, 115–122
6. Livett, E. A., Lee, J. A. and Tallis, J. H., *J. Ecol.*, 1979, **67**, 865–891.
7. Farmer, J. G., Swan, D. S. and Baxter, M. S., *Sci. Total Environ.*, 1980, **16**, 131–147.
8. Rippey, B., Murphy, R. J. and Kyle, S. W., *Environ. Sci. Technol.*, 1982, **16**, 23–30.
9. Batifal, F. M. and Boutron, C., *Atmos. Environ.*, 1984, **18**, 2507–2515.
10 Ruhling, A. and Tyler, G., *Water Air Soil Pollut.*, 1984, **22**, 173–180.
11. Thomas, M., Petit, D. and Lamberts, L., *Water Air Soil Pollut.*, 1984, **23**, 51–59.
12. Goldberg, E. D., Gamble, E., Griffin, J. J. and Koide, M., *Estuar. Coast. Mar. Sci.*, 1977, **5**, 549–561.
13. Bertine. K. K. and Goldberg, E. D., *Environ. Sci. Technol.*, 1977, **11**, 297–299.
14. Kemp. A. L. W., Williams, J. D. H., Thomas, R. L. and Gregory, M. L., *Water Air Soil Pollut.*, 1978, **10**, 381–402.
15. Nriagu, J. O., Kemp, A. L. W., Wong, H. K. T. and Harper, N., *Geochim. Cosmochim. Acta*, 1979, **43**, 247–258.
16. Wong, H. K. T., Nriagu, J. O. and Coker, R. D., *Chem. Geol.*, 1984, **44**, 187–201.
17. Greenberg, A. E., Trussell, R. R., Clesceri, L. S. and Franson, M. A. H., *Standard Methods for the Examination of Water and Waste Water*, APHA, AWWA, WPCF, Washington, 1985.
18. Ure, A. M. and Berrow, M. L., *The Elemental Constituents of Soils* (ed. Bowen, H. J. M.), *Environmental Chemistry*, The Royal Society of Chemistry, London, 1982, vol. 2, pp. 94–204.
19. Weast, R. C., *CRC Handbook of Chemistry and Physics*, Boca Raton, Florida, 1988.

ACKNOWLEDGEMENT. We thank Prof. P. S. Satsangi, Director, D. E. I. Dayalbagh, Agra, for moral support and encouragement.

Received 16 January 1995; revised accepted 12 May 1995

Mahatma Gandhi: An Apostle of Applied Human Ecology. T. N. Khoshoo. Tata Energy Research Institute, New Delhi. 1995. pp. 71.

1995 has been declared by UNESCO as the International Year For Tolerance. The principal aim of this initiative is to promote among humankind at least tolerance and understanding, if not love, of diversity and pluralism in human communities in terms of religion, language, ethnicity, skin colour, gender and political belief. The Director-General of UNESCO, Federico Mayor, mentioned in a lecture at New Delhi in May 1995 that fostering appreciation and tolerance of human diversity is the best way of commemorating the 125th birth anniversary of Mahatma Gandhi. Gandhiji's gospel of nonviolence was based on his conviction that nonviolence cannot be practised in compartments – it has to be total, encompassing all manifestations of life. Thus, he once mentioned that it will not be possible to be nonviolent to nature if we are going to be violent to each other. He was convinced that the alternative to *sarvodaya* was *sarvanasha*.

The Global Biodiversity Convention at Rio de Janeiro in 1992 designed to protecting flora, fauna and microorganisms will not help in achieving its purpose, unless there is an equal commitment to preserve, respect and nurture human diversity. It is worth repeating this central message of Gandhian thought and action at a time when violence is growing in the human heart. This trend is getting reflected in many ways – starting with violence to oneself (through drug addiction and AIDS) and violence to whole religious or ethnic groups as is happening in the former Yugoslavia or in Rwanda. The term 'ethnic cleansing' is now being used in the mass media without evoking horror and universal condemnation. T. N. Khoshoo has reminded us through this well-written book the beauty, simplicity and eternal relevance of Gandhiji's views on different aspects of human development. The book indicates how the four major ills of contemporary human developmental pathways could have been avoided had policy makers in all countries followed the Gandhian approach and methods of planning. It would be useful to refer to these briefly, since reexamining them and adapting them to present-day conditions could pave the way to achieving the harmony we now seek with nature and with each other.

The first serious malady of today's pattern of economic growth is the increasing rich–poor divide in economic well-being. Over 84% of the world's annual income now goes to 20% of the world's people. Another 20% get only 1.4% of the global annual income. One billion women and men live on a daily per capita income of less than one US dollar, while another 2 billion earn less than 2 US dollars per day. Everyday 68,000 babies are born to families living on less than one dollar per day. The Gandhian recipe to such gross inequity consisted of three seminal ideas. First, the promotion of sustainable life-styles would lead to curbing greed and meeting the essential needs of everyone. Second, the adoption of democratic decentralization in planning would lead to addressing the minimum needs of every child, woman and man. Third, the adoption of the concept of trusteeship by the well-to-do with reference to their surplus wealth will result in the promotion of symbiotic social contracts between the rich and the poor. For example, if the 20% of the global population now enjoying 84% of its annual income will regard themselves as trustees and not owners of their enormous income and will adopt sustainable life-styles and share their wealth with the poor, the funds needed to overcome extreme poverty and deprivation will be readily available.

Secondly, the present pathways of development are causing serious damage to the environment, resulting often in irreversible harm to the rights of the unborn for a healthy and productive life. Khoshoo has explained in detail the Gandhian approach to protecting our life support systems of land, water, flora, fauna, forests and the atmosphere. Conserving nature and natural resources and using them in an ecologically sustainable and socially equitable manner constitutes the core of the Gandhian approach to environment protection.

A third serious malady of present-day developmental paradigm is the phenomenon of jobless growth, i.e. GDP increases but employment opportunities diminish. Gandhiji's approach to promoting sustainable livelihood security was the creation of eco-jobs based on organic recycling, garbage and sewage utilization, generation of wealth from wastes, ecological farming, value–addition to natural products through village-level agro-processing and promotion of self-reliance at the village level to the extent possible. Decentralized production supported by a few key centralized services will help to convert the concept of sustainable livelihoods for all from rhetoric to reality. The *charkha* was the most elegant and effective symbol Gandhiji used to stress the need to convert cotton into cloth at the local and even household level, rather than sending cotton abroad (as was happening in the 1930s) and importing textiles. To the slogan 'export or perish' the Gandhian answer will be 'import less and live', since the present pattern of exports based on an unsustainable exploitation of environmental capital stocks would result in 'export and perish'.

Finally, the most serious of present-day ills is the spread of violence. Mini conflicts are replacing mega conflicts. Failure to integrate the principles of equity and ethics in the present developmental paradigm is a major reason for the spread of a culture of violence. Economic structural adjustment dictated by international financial institutions enhances the deprivation and misery of the poor. These adverse effects are being addressed through recipes such as structural adjustment with a human face and safety nets for the poor. Gandhiji considered economic equality an essential condition for nonviolent independence. He once mentioned that a technological society has two choices. First, it can wait until catastrophic failures expose systemic deficiencies, distortion and self-deceptions (as is happening now). Alternatively, a culture can provide social checks and balances to correct for systemic distortion prior to catastrophic failures. The alternative approach involves adopting *antyodaya* (attention to the poorest person) as a means to achieving *sarvodaya* (well-being of all). It also involves gender equity, with women being accorded the same status and attention as men in all spheres of human development. Thus, the structural adjustment we need now in the world is adjustment to sustainable life-styles and to the internalization of the principles of equity, both in gender and economic terms, in our daily life.

Khoshoo and the Tata Energy Research Institute have rendered great service to

contemporary development debate by bringing out this collection of Gandhian principles in the areas of environment and development. At least in vocabulary, the Gandhian approach is being recapitulated in the declarations of the recent UN Conferences. If a stripe review is undertaken of the declarations made at the Children's Summit in New York (1990), the UN Conference on Environment and Development, Rio de Janeiro (1992), the Conference on Human Rights, Vienna (1993), the Conference on Population and Development, Cairo (1994), the UN Social Summit, Copenhagen (1995) and the Fourth World Conference on Women to be held in Beijing in September 1995, we will find a striking resemblance between many of the statements contained in Khoshoo's book and those embodied in the international declarations. Converting the Gandhian development strategies and value systems from theory into practice is thus an idea whose time has come.

M. S. SWAMINATHAN

M. S. Swaminathan Research Foundation,
3rd Cross Street,
Taramani Institutional Area,
Madras 600 113, India.

Recent Trends in Aerobiology, Allergy and Immunology. A collection of Plenary Lectures and Contributory Articles. Ed. S. N. Agashe. Oxford & IBH Publishing Co. Pvt. Ltd., New Delhi. 1994. 315 pp. Price: Rs 750.

Aerobiology is the study of airborne particles of biological origin, their sources, liberation, deposition and impact on agriculture/human health. The presence of these particles in the atmosphere is influenced by environmental and climatic conditions. The emphasis of aerobiology has shifted from one aspect to another. During the last few decades, aerobiology has assumed importance because of its role in diagnosis of allergic disorders. The role of pollen in allergic rhinitis was shown by Charles Blackley as early as 1873. Since then, a lot of work has gone in to show the relationship between aerobiology and allergology. Various systematic studies were conducted on the presence of pollens and fungal spores in the atmosphere and their effect on human health. The book under review consists mainly of papers presented at the 5th International Aerobiology Conference held at Bangalore in August 1994. The 20 topics cover diverse areas, mainly of aerobiology related to allergy, except for a few topics dealing with the usefulness of aerobiology in forensic sciences, solar radiation, polymorphic light eruption, and paradox and placebo of medicine in the 21st century.

Brown in his article has brought out the historical relationship between aerobiology and allergy, while John Lacey in his article has emphasized the other aspects of aerobiology and given various priorities for aerobiological research. Other articles on pollen calendar and the environmental influences on the allergenicity of pollens give a lot of information in this area. A few articles on fungal spores, including alternaria, and their effect on human health and foods are worth reading. The pollen allergy in India has been dealt with in great detail. Also there is an interesting article on latex allergy which warns of the severe risk for workers in the manufacture of latex products. The house dust mite is a known source of allergen which has been shown to be a principal factor responsible for early morning asthmatic attacks. The book has two articles on this topic: one on the current studies of research on dust mite in India and the other on the use of mite allergens for immunotherapy. Until a few years ago, insects were not considered as major allergenic source. However, during the last few years insects as aeroallergens have been conclusively established and the review of this topic is very informative. The book also covers molecular characterization of allergens and the current studies in immunotherapy.

The wide range of fields covered in this volume illustrates the importance of aerobiology. However, the topics are not arranged with any specific aspect in mind. There is scope for improvement by reducing the typographical errors and improving the quality of paper and the production of photograhs. Even though the references quoted by various authors are comprehensive, there are serious deficiencies concerning recently published literature. The inclusion of subject index and abbreviations used in the book could have added to the quality of the book.

In spite of all deficiencies the book gives a lot of current information on various aspects of aerobiology and is an important edition for the students working in this field. The book is not priced but I hope it will be within the reach of any Indian student and will be inspiring to the research workers in aerobiology.

S. V. GANGAL

Centre for Biochemical Technology,
Delhi University Campus,
Mall Road, Near Jubilee Hall,
Delhi 110 007, India.

***In vitro* Culture and its Applications in Horticulture.** H. Vidalie (coordinator). Oxford & IBH Publishing Co. Pvt. Ltd., New Delhi. 1995. 231 pp. Price: Rs 360.

In vitro culture of economically important crop species is a key area of applied research in biotechnology. Micropropagation technology is the best example of direct applications of research outcome in different plant species. During the past decade many '*in vitro*' culture techniques have been scaled up to commercial level and now various commercial companies are producing a large number of propagules of selected rare clones of various plant species, among them ornamental and horticultural crops being the main candidates. In these species full-fledged technologies are standardized and are being commercially exploited. A lot of information is available in the form of research and review articles. There are many books available on the *in vitro* culture of ornamental, horticulture, woody perennials, and important crop species. The book under review is an addition to the voluminous literature already existing. The book is contributed by a team of teachers, researchers and practitioners of commercial tissue culture from Angers, France.

This book covers the various aspects of *in vitro* culture application, main methods of laboratory set-up, media preparation, the choice of explant, and physiological phenomena related to *in vitro* culture. The genetic aspect of tissue-culture-raised plants has also been considered.

The first part of the book is on the history, basic physiology of differentiation and redifferentiation, the role of hormones, the choice of explant as per age, size, position and response in culture. One of the chapters is on nutritional aspects, such as chemical composition of the medium and its physical characteristics (e.g. liquid/solid environment, media and gaseous exchange). Physical characteristics of the culture environment such as the intensity, quality and duration of light, and the temperature are also discussed.

The chapter on the technology of *in vitro* culture includes information on laboratory set-up and instruments/equipment for *in vitro* culture, the complete laboratory procedures for preparation of glassware, stock solutions, media and their sterilization. Procedures for the preparation, disinfections are also discussed.

Meristem culture is illustrated with 'Japanese artichoke' as a model system and also micropropagation, giving an example of *Saintapaulia*. Also covered is the general background of techniques and significance of anther/pollen culture and protoplast culture.

The chapter 'Biotechnologies in horticulture: Possibilities and perspectives' describes modern methods such as protoplast fusion, haploids, somaclonal variation, and molecular genetics. There is also a chapter on gene transformation and somatic embryogenesis, and on genetic aspects of *in vitro* regenerated plants.

The book contains an informative article on the mycorrhiza and *in vitro* culture, where the types of mycorrhiza, its association with specific plants, and applications of mycorrhization have been explained in detail. There is one chapter which deals with application of *in vitro* culture in disease elimination and disease resistance.

Lastly, advantages of *in vitro* culture, its economical aspects and constraints are presented. There is an account of the present and future status of *in vitro* culture of ornamental plants, woody plants and fruit trees. A list of plants for which technology has been commercialized is given.

Overall, the book covers information on general techniques of *in vitro* culture useful for beginners and students. The information could have been better organized and repetition of many aspects avoided. Although recent applications and constraints of *in vitro* culture have been covered, they should have been described in more detail so as to make the information more complete. As the title of the book indicates, an account of *in vitro* technologies for economically important horticultural crops and significance of *in vitro* culture pertaining to specific problems in these crops would have been more useful.

Tissue culture and its commercial exploitation are emerging fields and this book will be useful to people entering these fields.

R. S. NADGAUDA

Plant Tissue Culture Division,
National Chemical Laboratory,
Pune 411 008, India.

Erratum

Testosterone metabolism by rhesus monkey spermatozoa: Effects of antifertility agents

M. Rajalakshmi, R. S. Sharma, S. Mokkapati and P. C. Pal
(*Curr. Sci.*, 1995, **68**, 1121)

The title of the article and the names of the authors were inadvertently left out from the Contents of the 10 June issue. We regret the error.

— *Editors*

Two undocumented historic earthquakes and additional information on the major earthquake of 1 September 1803 from north India

R. N. Athavale

Information about historic earthquakes noted in the records of the Peshwa period is extracted. It reveals in north Indian plains occurrence of two earthquakes not documented in the existing catalogues of Indian earthquakes. One event was felt near Delhi on 10 October 1760 and the other at Varanasi on 30 September 1803. Some additional information about the documented earthquake of 1 September 1803 has become available. It suggests that this major earthquake had its epicentre in the Garhwal area of Himalaya and it caused simultaneous damage in Allahabad, Varanasi and Gaya cities in the Gangetic plains.

The recent earthquakes of Latur on 30 September 1993, Uttarkashi on 20 October 1991 and Koyna on 11 December 1967 and the current discussion on advisability of the construction of Tehri dam in an earthquake-prone area have increased public awareness on the need of seismic studies and surveillance. Catalogues of historic and recent earthquakes have been prepared by various workers from time to time[1–10]. These inventories of earthquakes are useful in knowing about general seismicity level of an area and in preparing seismic zoning maps.

The Koyna earthquake had an initial element of surprise as it took place in an area generally considered by earth scientists as aseismic. Subsequent to the Koyna earthquake, Kelkar[11] searched the historical records of Maharashtra over the last 300 years and showed that the Western Continental margin in general and the Koyna area in particular were not as aseismic as was believed till then. He produced historical evidence for ten earthquakes over the period 1598–1832. Out of these, two earthquakes of 1757 and 1764 AD were in the Dhom–Mahabaleshwar area, where the dam on Koyna river is located.

I have recently come across another article in the 26 January 1934 issue of the Marathi daily *Kesari*[12] which contains description of two earthquake events of 10 October 1760 and 30 September 1803 which are not listed in the compilations referred to above. Further, it also provides a more graphic description about the listed major earthquake of 1 September 1803, which affected large areas of north India. The following description is about the 1 September 1803 (Bhadrapad, Shuddha Chaturdashi) earthquake:

'The earthquake took place at midday. Houses oscillated for one Ghataka (24 min). Persons sleeping on the side were turned either face upwards or downwards. The cattle on leash showed great unease. Roof tiles fell. The grain market area caught fire. Two stars fell. Their light during daytime was similar to starlight and lasted for a quarter Ghataka (6 min). People got scared. They left their houses and came out in courtyards. Everyone felt as if doomsday was today itself. This earthquake occurred at Kashi (Varanasi), Prayag (Allahabad) and Gaya, at these three places at the same time. Besides, up above near Lord Vishnu's abode Badrikedar, the city of Srinagar existed at a distance of 4–5 kos (10–15 km); in that city and in its neighbourhood, earthquakes occurred. The land parted. About 500 to 1000 houses collapsed; mountains also collapsed'.

The reference to Srinagar town of Garhwal district of Uttar Pradesh in past tense in the above account is significant. Baird-Smith[1] states that Srinagar, the capital of Garhwal state, had 700–800 houses in 1796 when Colonel Hardwicke visited it and that it was found in a ruined status by M/s Webb and Raper when they visited it in 1808. The information culled from Peshwa period records about damage to Srinagar town is thus corroborated through independent observations. Baird-Smith[1] reports on the damage caused by this earthquake in Mathura and Oldham[2] records that it was felt in Calcutta and also in the upper Ganga valley. The same earthquake was felt at Lucknow, Farrukhabad and Jabalpur[13] also (Figure 1).

Chandra[6] lists the Mathura shock and the one felt in the upper Ganga valley as separate events of 1 September 1803 and does not include the Calcutta shock for this date. Bapat *et al.*[3] list only the Mathura shock against 1 September 1803 and assign it a magnitude of 6.5. Rastogi[14] assigns a magnitude of 7 to this earthquake. Gupta[15] and Seeber and Ambruster[16] have considered four definite great earthquakes of Indian Himalaya as the 1897 Shillong, 1905 Kangra, 1934 Bihar–Nepal and 1950 Assam events and three doubtful ones as the 1803 shock discussed here, the 1833 even in western Bihar and the 1885 shock of Kashmir. The additional historic information gleaned from the Peshwa period record suggests that the earthquake of 1 September 1803, with its meizoseismal area in the Garhwal region, affecting the entire Gangetic alluvial plain from Mathura to Calcutta and felt at Jabalpur (Figure 1) in the shield area, can also be included in the category of great earthquakes (magnitude greater than 8) without any reservation.

The descriptions[12] of the other two earthquakes which do not find a place in the earthquakes catalogues referred to above are as follows.

Regarding the earthquake of 30 September 1803:

'On Ashwin Shuddh Chaturdashi (30 September 1803) two big stars fell in Kashi (Varanasi). Light shined for 4 Ghatakas (1 h 36 min). There was loud sound similar to that made by a thunderbolt. A big earthquake took place. Two-three hundred houses fell. The Bhagirathi was very turbulent for 4 Ghatakas. Rainfall occurred for one-and-a-half day. In the area (Moholla) near Bindumadhav, up to 200 houses got burnt. The water in the stream of Ganga was turbulently rising up to two spears height for a period of 3 Ghatakas (1 h 12 min). And many loud sounds occurred in the water and thousands of stars (Nakshtras) fell at night.'

The second earthquake is reported by Nana Phadnavis, a minister at Peshwa's court, who accompanied the Maratha army

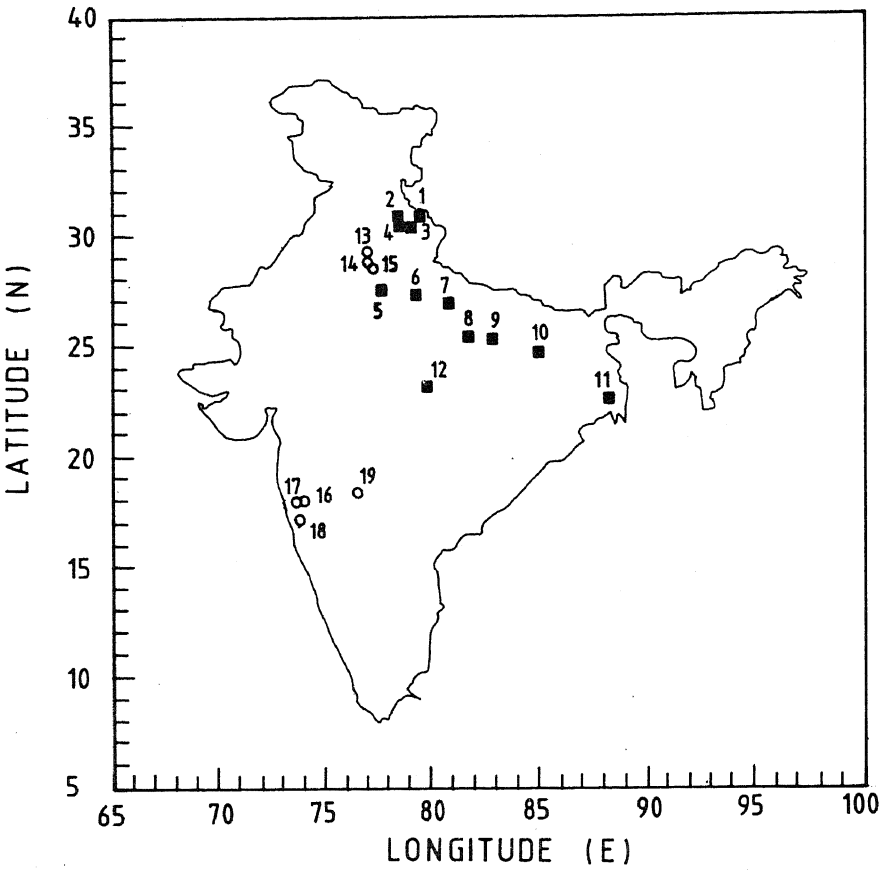


Figure 1. Map of India showing towns/sites described in the text. Square symbols are for localities affected by the earthquake of 1 September 1803. [1. Badrinath, 2. Uttarkashi, 3. Srinagar, 4. Tehri, 5. Mathura, 6. Farrukhabad, 7. Lucknow, 8. Allahabad, 9. Varanasi, 10. Gaya, 11. Calcutta, 12. Jabalpur, 13. Panipat, 14. Sonipat, 15. Delhi, 16. Dhom, 17. Mahabaleshwar, 18. Koynanagar, 19. Latur.]

on its northern campaign which culminated in the third battle of Panipat[17] fought on 14 January 1761. He records the following in his autobiography:

'There I worshiped (darshan) the Lord of the Earth (Prithivipati) as ordered by his Highness (Shriman) Bhausaheb (Peshwa). Then i came to His Highness at his army camp. Afterwards, an earthquake occurred while I was sitting with His Highness.'

The Maratha army was in Delhi–Panipat area from August 1760. According to historians, the ruler of Delhi (Badshah) was referred to in the contemporary Marathi literature as Prithvipati. Nana Phadnavis attended the anointment ceremony of Crown Prince Jawanbakht at Delhi on 10 October 1760 and then reached the Maratha army camp near Sonipat, about 45 km north of Delhi, on the same or perhaps on the next day. He, therefore, felt the earthquake at this location.

The information provided here and in Kelkar's article[11] suggests that our historical records are replete with descriptions pertaining to earthquakes and other catastrophic environmental events such as floods, droughts and meteorite falls. There is clearly a need for a forum for discussions between scientists interested in neotectonics, archaeoseismology and the past global changes (PAGES) programme of IGBP in India on the one hand and the historians and archaeologists on the other, so that the latter can look at their source material with the perspective of the former group. It also follows that any feasibility study for a major irrigation, industry or power project should incorporate a report on the environmental history of the project site, prepared by those who have worked on the history and archaeology of the area where the project is to be located.

1. Baird-Smith, R., *J. Asiatic Soc. Bengal*, 1843, **12**, 1029–1056.
2. Oldham, R., *Mem. Geol. Surv. India*, 1883, **19(3)**, 163–215.
3. Bapat, A., Kulkarni, R. C. and Guha, S., K., *Indian Soc. Earthquake Technol.*, 1983, 1–211.
4. Rao, B. R. and Rao, P. S., *Bull. Seismol. Soc. Am.*, 1984, **74**, 2519–2533.
5. Chandra, Umesh, *Bull. Seismol. Soc. Am.*, 1977, **67**, 1387–1413.
6. Chandra, Umesh, *Curr. Sci.*, 1992, **67**, 40–71.
7. Srivastava, H. N. and Ramachandran, K., *Mausam*, 1985, **36**, 351–358.
8. Ramachandran, K. and Srivastava, H. N., *Mausam*, 1991, **42**, 171–182.
9. Valdiya, K. S., *Curr. Sci.*, 1992, **63**, 291–296.
10. Guha, S. K. and Basu, P. C., 1993, AERB Tech. Doc. No. TD/CSE-1, pp. 1–70.
11. Kelkar, Y. N., *Kesari* (Marathi daily published from Pune), 7 January 1968, p. 7.
12. Anonymous, *Kesari* (Marathi daily published from Pune), 26 January 1934.
13. Chandra, Moti, *History of Kashi* (in Hindi), Hindi Granth Ratnakar, Bombay, 1962, p. 366.
14. Rastogi, B. K., *Mem. Geol. Soc. India*, 1995, **30**, 43–50.
15. Gupta, H. K., *Curr. Sci.*, 1992, **62**, 257–263.
16. Seeber, L. and Ambruster, J. G., in *Earthquake Prediction – An International Review*, Maurice Ewing Series 4, Am. Geophys. Union, Washington, D.C., 1981, pp. 259–277.
17. Joshi, L. (ed.), *Marathi Vishwa Kosh (Encyclopaedia)*, 1980, **9**, 547–550.

ACKNOWLEDGEMENTS. Mr K. G. Kulkarni has rendered valuable help in fixing the date and location of the 1760 earthquake. Mr Y. S. Lele invited my attention to the article in the daily *Kesari*. Prof. K. D. Abhyankar verified the conversion of 'tithis' in Hindu Almanac into Gregorian calendar dates. Discussions with many colleagues were useful.

R. N. Athavale is in the National Geophysical Research Institute, Hyderabad 500 007, India.

Edited and published by Prof. S. Ramaseshan and Prof. P. Balaram, Current Science Association, Bangalore 560 080.
Typeset by R D Typesetters, Bangalore 560 010. Printed by Lotus Printers, Bangalore (Ph: 3380167).

CURRENT SCIENCE

A fortnightly journal of research

Editors: S. Ramaseshan and P. Balaram

Editorial Board

S. Arunachalam, Central Electrochemical Research Institute, Karaikudi 623 006
Asis Datta, School of Life Sciences, Jawaharlal Nehru University, New Delhi 110 067
D. Balasubramanian, Centre for Cellular and Molecular Biology, Hyderabad 500 007
J. Chandrasekhar, Department of Organic Chemistry, Indian Institute of Science, Bangalore 560 012
R. Gadagkar, Centre for Ecological Sciences, Indian Institute of Science, Bangalore 560 012
K. N. Ganeshaiah, Department of Plant Genetics and Breeding, University of Agricultural Sciences, Bangalore 560 065
R. V. Hosur, Tata Institute of Fundamental Research, Bombay 400 005
N. V. Joshi, Centre for Ecological Sciences, Indian Institute of Science, Bangalore 560 012
Jyotirmoy Das, Indian Institute of Chemical Biology, Calcutta 700 032
C. C. Kartha, Division of Cellular and Molecular Cardiology, Sree Chitra Tirunal Institute for Medical Sciences and Technology, Thiruvananthapuram 695 011
S. S. Krishnamurthy, Inorganic and Physical Chemistry Department, Indian Institute of Science, Bangalore 560 012
S. C. Lakhotia, Department of Zoology, Banaras Hindu University, Varanasi 221 005
K. Muralidhar, Department of Zoology, Delhi University, Delhi 110 007
R. Narasimhan, CMC Limited, Bangalore 560 001
K. Neelakantan, ANURAG (Advanced Numerical Research and Analysis Group), Hyderabad 500 258
R. Nityananda, Raman Research Institute, Bangalore 560 080
T. Padmanabhan, Inter-University Centre for Astronomy and Astrophysics, Pune 411 007
B. R. Pai, Propulsion Division, National Aerospace Laboratories, Bangalore 560 017
T. J. Pandian, School of Biological Sciences, Madurai Kamaraj University, Madurai 625 021
Y. S. Rajan, Department of Science and Technology, New Delhi 110 016
M. Ramakrishnan, Geological Survey of India, Calcutta 700 016
A. V. Ramani, T. T. K. Pharma Ltd., Bangalore 560 025
S. Ramaswamy, Department of Physics, Indian Institute of Science, Bangalore 560 012
A. K. Raychaudhuri, Department of Physics, Indian Institute of Science, Bangalore 560 012
S. R. Shetye, Physical Oceanography Division, National Institute of Oceanography, Dona Paula, Goa 403 004
V. Siddhartha, 51 Bharati Nagar, New Delhi 110 003
V. Srinivas, School of Mathematics, Tata Institute of Fundamental Research, Bombay 400 005
G. Srinivasan, Raman Research Institute, Bangalore 560 080
R. Uma Shaanker, Department of Crop Physiology, University of Agricultural Sciences, Bangalore 560 065
K. S. Valdiya, Department of Geology, Kumaun University, Nainital 263 002
G. Venkataraman, Department of Physics, Sri Sathya Sai Institute of Higher Learning, Prasanthi Nilayam 515 134
M. Vidyasagar, Centre for Artificial Intelligence and Robotics, Bangalore 560 001
K. VijayRaghavan, TIFR Centre, Indian Institute of Science Campus, Bangalore 560 012

Current Science, founded in 1932, is published by the Current Science Association in collaboration with the Indian Academy of Sciences. The journal is also intended as a medium for communication and discussion of important issues that concern science and scientific activity. All articles published in *Current Science*, especially editorials, opinions and commentaries, letters and book reviews, are deemed to reflect the individual views of the authors and not the official points of view, either of the Current Science Association or of the Indian Academy of Sciences.

Working Committee, Current Science Association

P. Balaram, Indian Institute of Science, Bangalore
R. Chidambaram, Department of Atomic Energy, Bombay
V. K. Gaur, CSIR Centre for Mathematical Modelling and Computer Simulation, Bangalore
K. Kasturirangan, ISRO Satellite Centre, Bangalore
P. K. Kaw, Institute for Plasma Research, Gandhinagar
N. Kumar, Raman Research Institute, Bangalore (*Vice-President*)
N. V. Madhusudana, Raman Research Institute, Bangalore (*Secretary*)
A. Mani, Field Research Unit, Indian Institute of Tropical Meteorology, Bangalore (*Past-President*)
R. A. Mashelkar, National Chemical Laboratory, Pune
R. Narasimha, Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore
G. Padmanaban, Indian Institute of Science, Bangalore (*Vice-President*)
R. Ramachandran, The Institute of Mathematical Sciences, Madras
P. Rama Rao, Department of Science and Technology, New Delhi
S. Ramaseshan, Raman Research Institute, Bangalore (*Treasurer*)
M. S. Valiathan, Manipal Academy of Higher Education, Manipal (*President*)

Editorial Office

Editorial staff

Chandrika Ramesh, G. Madhavan, C. S. Ravi Kumar, M. S. Venugopal

Circulation and Accounts

Peter Jayaraj, Ranjini Mohan, B. Sethumani, Shanthi Bhaskar, B. K. Shivaramiah

Annual subscription 1995

	Institutional	Personal
India	Rs 350	Rs 100
All other countries (AIR MAIL)	US $175	US $50

Single copies other than special issues: Rs 35/US $15

Editorial Office: CURRENT SCIENCE
C. V. Raman Avenue
P.B. 8001
Bangalore 560 080, India
Telephone: 91-80-3342310
Telex 845-2178 ACAD IN
Fax: 91-80-3346094

Copyright © 1995, Indian Academy of Sciences

NEW ICAR AWARDS

The Indian Council of Agricultural Research, New Delhi announces the constitution of the following New Awards to be given away from 1995 onwards.

1. **Best Institution Award**

   In order to recognize the best performance in Agricultural Research, Extension and Education, three Agricultural Awards of Rs 100,000/- each will be given to two ICAR institutes/NRC/Project Directorates and one State Agricultural University, partially out of the interest accrued from the King Boudin Award bestowed on the ICAR in 1989.

2. **Best KVK Award**

   In order to recognize the best performance of Krishi Vigyan Kendras, KVKs which have run for a minimum period of 5 years, two awards of Rs 50,000/- each will be given in alternate years. The performance will be judged in terms of vocational training, on-farm research training of the extension workers and other extension programmes, keeping in view the mandate of the KVKs.

3. **Best Extension Workers Award**

   To create incentives for best extension workers in developing extension education programmes, adopting extension methodologies and creative impact of the programme on the farming community, especially on the downtrodden, four awards of Rs 10,000/- each are to be awarded once in three years. The first award will be given in 1996 for the preceding three years i.e. 1993–1995.

   The application forms for the above awards may be obtained from

   Dr R. C. Maheshwari, Assistant Director General (CSC), ICAR, Krishi Bhavan, New Delhi, on or before 26 September 1995. The applications may be forwarded by the Head of the Institution/Directors of Institutes/State Department of Agriculture/ Horticulture/Animal Science/Fisheries and Vice-Chancellors of State Agricultural Universities. The last date for submission of applications for these awards is 1 November 1995. The last date for candidates in the Andaman and Nicobar Islands, Lakshadweep, States, Union Territories in the North Eastern Region, Ladakh Division of J&K State and Sikkim is 15 November 1995.

   Note: A copy of the proforma can be obtained from him by sending a self-addressed one-rupee stamped envelope.

Founded: 1932

ISSN: 0011-3891

CURRENT SCIENCE

Volume 69 Number 4

25 August 1995

283 **In this issue**

NEWS

285 ***Chemical Abstracts* on CD-ROM**
Subbiah Arunachalam

285 **Intensive course on 'applications of remotely sensed data on ocean colour'**
P. S. Swathi

286 **International Lithosphere Project**
Ramesh P. Singh

CORRESPONDENCE

287 **Necessity of broadening science base in India,** Mariappan Periasamy ■ **287 More on the NSU,** N. S. Narasimhan ■ **288 Response,** P. N. Srivastava

SCIENTIFIC CORRESPONDENCE

288 **Comments on 'What kind of a system is a human being? A cybernetic system?' [*Curr. Sci.*, 1994, 67, 936–940] ■ 289 Comments on the review of the monograph *The Scientific Methodology in the Light of Cybernetics* [*Curr. Sci.*, 1994, 67, 954–956],** P. R. Masani; ■ **292 *Reply by*,** R. Narasimhan

293 **Methane emission from rice paddies: Need for a downward revision of global estimate**
J. S. Singh and Smita Singh

295 **Nomenclature of super heavy elements**
S. K. Arya and S. S. Vasan

SPECIAL SECTION: ANIMAL WELFARE

296 **Whose life is it anyway? A review of some ideas on the issue of animal rights**
Anindya Sinha

301 **On the use of animals in research and education**
Jane Goodall

303 **Human–animal interactions**
N. K. Waran

305 **Conflicts or coexistence?**
Madhusudan Katti

308 **Environmental design**
N. K. Waran

310 **Stereotypies in the stabled horse: Causes, treatments and prevention**
L. Winskill, N. K. Waran, C. Channing and R. Young

GENERAL ARTICLE

317 **Seamless chemical engineering science: The emerging paradigm**
R. A. Mashelkar

REVIEW ARTICLE

336 **The neurotransmitter receptor triad: Nicotinic acetylcholine receptor, γ-aminobutyric acid receptor (type A) and inhibitory glycine receptor**
P. R. Satish and A. S. Balasubramanian

RESEARCH ARTICLES

343 **On the usefulness of quantum chemical bond order as a local reaction path indicator: The case study of a model carbonyl addition reaction**
D. K. Maity, D. Majumdar and S. P. Bhattacharyya

351 **High-temperature tolerance of *Petunia* and *Nicotiana* pollen**
G. U. Rao, K. R. Shivanna and V. K. Sawhney

RESEARCH COMMUNICATIONS

356 **Enhanced giant magnetoresistance in magnetic/nonmagnetic $La_{1-x}MnO_{3-\delta}$ oxide multilayers**
S. Sundar Manoharan, K. M. Satyalakshmi, V. Prasad, S. V. Subramanyam and M. S. Hegde

358 **Physiobiochemical aspects of development of hardseededness in *Albizia lebbek***
R. C. Uniyal and A. R. Nautiyal

361 **Fractal description of seismicity of India and inferences regarding earthquake hazard**
K. N. Khattri

BOOK REVIEWS

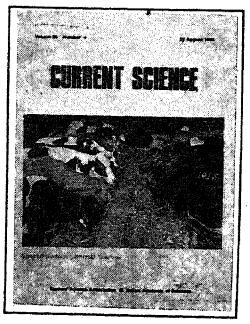
367 **Beyond the Quantum Paradox, *reviewed by* Habeebul Haque Ansari ■ 368 The Action Principle in Physics, *reviewed by* Deepak Dhar ■ 369 Laboratory Manual of Organic Chemistry, *reviewed by* R. Balaji Rao ■ 370 Perspectives in Nematode Physiology and Biochemistry, *reviewed by* M. Shamim Jairajpuri ■ 370 Ocean Technology – Perspectives, *reviewed by* D. Srinivasan ■ 371 Botany in India: History and Progress, *reviewed by* D. D. Awasthi**

PERSONAL NEWS

373 **Prof. T. R. Govindachari – An appreciation**
K. Nagarajan

HISTORICAL COMMENTARY AND NOTES

375 **Eugene Paul Wigner – A tribute**
N. Mukunda



COVER. Exploitation of animals has been mostly promoted by a general lack of concern for their welfare associated with our meaningless value systems. A series of articles in this issue (pages 296–316) explores these issues.(The cover photograph and several in the special section are by K. Chandrapal Singh.)

Indexed in CURRENT CONTENTS/GEOBASE/CHEMICAL ABSTRACTS

The editors thank Dr. K. N. Ganeshaiah, Department of Plant Genetics and Breeding, University of Agricultural Sciences, GKVK, Bangalore 560 065, and Dr. Anindya Sinha, National Centre for Biological Sciences, TIFR Centre, P.O. Box 1234, Bangalore 560 012, for agreeing to be guest editors.

In this issue

Tracking transition states through bond orders

The most direct approach to the understanding of reaction mechanisms is by examining potential energy surfaces in detail. A few stationary points, viz. the minima characterizing the reactants, products and potential intermediates and saddle points connecting them in an idealized cross-section, are specially important. Within transition state theory, the information about these points is just about everything one may wish to know about the course of the reaction. Minima are directly accessible to experimental investigations. But transition states are elusive. Kinetic experiments can yield activation energies, but not any precise details of other properties of the transition state.

Chemists are not put off by the difficulties associated with precise characterization of transition state structures. They invariably visualize reactions in terms of making, breaking or rearrangement of a few bonds. The 'action' is assumed to be localized in a small fragment even in a complex substrate. The simplified notions serve their purpose; after all, many reactions are used in predictable ways to make new compounds. But can the 'arrowpushing' approach be placed on a firmer footing? Can theory help?

Bond order, in spite of being only a computable property, is a concept which has found widespread acceptance among chemists. It is intuitively reasonable besides being intimately related to measurable quantities like bond lengths and bond energies. The latter are relatively difficult to calculate using quantum chemical methods for large molecules. On the other hand, sensible bond orders can be obtained even with fairly simple computational methods. Variations in bond orders have been suggested to be useful indicators of the changes occurring during the course of a description for many simple reactions in the last few years. The inflexions noted in the bond orders have also been suggested to correlate well with the location of transition states. A detailed analysis with more complex examples is now provided (Maity *et al.*, **page 343**).

The approach employed by the authors is not going to replace conventional interpretations based on potential energy surfaces. The work is not meant in that spirit either. Just as orbital energy changes along the reaction coordinate can provide additional insights when interpreted with care, the methodology proposed based on bond orders represents another framework for understanding reactions. The language would also be familiar to nontheoreticians.

Animal welfare

Shivi, a king in the Mahabharatha, once gave protection to a dove which was being pursued by a hawk. The hawk demanded that the dove be handed over, but Shivi refused to do so, saying that he was obliged to protect it. The hawk then protested that this would mean starvation for him, and that Shivi himself would be responsible for the death of yet another living being in his attempt to save one life. In response, Shivi drew out his sword and proceeded to cut off his own flesh to offer it to the hawk and at the same time save the dove. The conflict portrayed here between the kind act of the king to save the dove and the cruelty that it would otherwise inflict on the hawk is almost eternal and haunts us in our ways of dealing with animals. The lesson conveyed, however, is simple and effective: if we do have the will, a little sacrifice on our part can achieve a lot for the welfare of animals.

King Shivis we may not be, but we could at least occasionally spare a thought for our fellow animals, exploited mercilessly ever since the dawn of our 'civilization' so that we may live a little longer, a little better. This is the theme of the special section on animal welfare in this issue of *Current Science*, a series of articles which attempt to present various aspects of our use of non-human animals.

The way we treat animals might largely rest on the contemporary philosophical strength of our civilization and on the extent of our realization that we, the humans, have a duty to think and worry about the welfare of animals just as we do about ourselves. Anindya Sinha (**page 296**) writes about these duties of the human race and reviews some of the old and new ideas that have guided our ethics of animal exploitation. This is followed by an appeal (**page 301**) by the noted primatologist Jane Goodall to initiate a dialogue on how noninvasively and painlessly animals can be used, if at all they must be, in the course of experimentation and education. She also identifies the responsibilities of the institutions of higher learning and the role they can play in minimizing the ill treatment that animals are often subjected to.

The tender relationship the farmers have with their herds of cattle, goat and sheep in Indian villages probably forms an exemplary way to incorporate the lesson from Shivi's legend. These farmers often take care of their domestic animals with extreme kindness, the kind that they show only to their own children, but at the same time they also subject them to hard work for their livelihood. It is not unusual to find, in our villages, the farmers tenderly touching their animals, hand-feeding the cattle late at night, and grooming them as they are given rest amidst the heavy working schedule. Does this physical touch elicit any healthy

response from the cattle? Natalie Waran in her article (**page 303**) on human–animal interactions explores the evidence that these do indeed greatly affect the psychological disposition of animals towards humans, their own survival, and possibly even productivity. Unfortunately, following the introduction of the western mechanized system of dairy management in India, the way cattle are treated has been changing: from a family companion to a mere productive unit. This dichotomy between the east and the west in their treatment of animals is the basis of the article by Madhusudan Katti (**page 305**), who outlines the actual realities that characterize human–animal conflict in most situations.

Natalie Waran, in yet another article (**page 308**), discusses various modifications of the immediate environment of the captive animal which should be achieved in response to its physical, social and developmental needs. The importance of both these approaches for the well-being of our otherwise exploited livestock cannot be overemphasized. Finally, N. Winskill *et al.* draw attention to a specific example of the devastation that our maltreatment of domesticated animals can wreak (**page 310**), with a paper on the behavioural abnormalities of stabled horses.

In today's difficult world, it is often quite easy to ignore whatever does not attract our direct attention. We may thus never give a thought to the hundreds of cows that produce our breakfast milk, nor remember the thousands of monkeys that have been sacrificed to produce one effective drug. At the other end of the spectrum, it may be impossible to change the ways in which the majority of our self-driven humanity thinks. There are, however, small steps that we can take so that the animals that we use can live life somewhat better, possibly a little more painlessly. As Edmund Burke said: '... nobody makes a greater mistake than he who did nothing because he could only do a little ...'.

K. N. GANESHAIAH
ANINDYA SINHA

More on the NSU

The debate on the National Science University was closed with the publication of the final reports on the NSU in the 25 June issue of *Current Science*. However, we have received more correspondence, some of which appears in this issue. Future letters on this subject will be published only if new points that have not been stressed so far are made by correspondents.

P. B.

CURRENT SCIENCE

Volume 69 Number 4 25 August 1995

NEWS

Chemical Abstracts on CD-ROM

Chemists around the world will welcome the CD-ROM version of their favourite search tool, *Chemical Abstracts (CA)*, described by its publishers as 'the key to the chemical literature', when it becomes available in 1996.

CA on CD-ROM will provide easy access to the same information available in the printed version–complete bibliographic citations, abstracts, structure diagrams, keyword indexing and comprehensive volume indexing–but with the added advantages of the electronic medium. *CA* currently covers about 600,000 documents every year, including articles from over 8000 journals. The new product will be compatible with both Windows and Macintosh systems and a user can access information from a number of access points, such as author name, keyword, substance name and chemical formula. On the price front, though, customers will have to pay an as-yet-undisclosed additional sum over the subscription price of the printed version. In return, a basic subscription will allow four simultaneous users at the same site.

Some time ago, Chemical Abstracts Service (CAS) brought out the compact disc version of the Twelfth Collective Index of *CA*, but users around the world found it less than ideal. Hope fully, the search software of *CA* on CD meets the expectation of the users.

CAS will do well to bring out *CA* of at least the past ten years on CD. Chemists tired of handling the heavy bound volumes of *CA* will be grateful. Indeed, *CA* is perhaps the last of the major science databases to be issued on CD-ROM format. The citation indexes of ISI, Philadelphia, BIOSIS, Compendex Plus and INSPEC's *Physics Abstracts* and related products are available on CD for a number of years.

Subbiah Arunachalam, Central Electrochemical Research Institute, Karaikudi.

Intensive course on 'applications of remotely sensed data on ocean colour'*

Primary production in the ocean lies at the base of the marine food chain. An accurate estimation of primary production has, therefore, several important applications: (a) it gives an indication of the amount of food available to sustain higher forms of marine life; (b) it plays a significant role in the thermodynamics of the mixed layer, which, in turn, controls the exchange of energy between the ocean and the overlying atmosphere and; (c) it provides a lower bound on the carbon dioxide flux across the air–sea interface. Fortunately, satellite imagery is providing ever more reliable estimates of primary production. This comes as a most welcome development since the spatial and temporal coverages of conventional measurements are extremely sparse.

Fifty scientists, including 22 foreign participants from 18 different countries, participated in the course.

Shubha Satyendranath, Dalhousie University, and Trevor Platt, Bedford Institute of Oceanography, both eminent researchers in this area, were instructors for the course. The topics covered included a hierarchy of models and algorithms of varying complexity to estimate primary production (notably, nonspectral, spectral, available light and absorbed light models), determination of total water column productivity, extrapolation of local algorithms to regional and basin scales and the use of ocean colour as a tool in the modelling of mixed-layer processess and biogeochemistry of the ocean. There were also five invited talks by participants on instrumentation and sensors and six on the application of ocean colour data.

The organization of the course was particularly timely, as the imminent launch of the SeaWIFS and other ocean colour satellites would soon provide considerable amount of data, which, in turn, would call for extensive modelling and analysis. The comprehensive coverage of all aspects of the analysis of ocean colour data and the sharp focus of the course was, therefore, appropriate.

P. S. Swathi, CSIR Centre for Mathematical Modelling and Computer Simulation, Bangalore.

*Sponsored by The Department of Ocean Development along with 7 international agencies (COSTED, IOC, JGOFS, ROSTAS, SAREC, SCOR, START and UNESCO) and held at the CSIR Centre for Mathematical Modelling and Computer Simulation (C-MMACS), Bangalore, during 3–7 April 1995.

International Lithosphere Project

An organizational meeting of the Indian section of the ILP task group II-4 and a seminar on 'Three-dimensional mapping of the lithosphere' was held at the Indian Institute of Technology, Kanpur, on 30 March to 1 April 1995. The meeting and the seminar were attended by scientists from a number of research and academic institutions in India. In the inaugural lecture, Vinod Gaur (CMMACS, Bangalore) discussed the structure and kinematics of the Indian continental lithosphere, including his recent work on GPS-derived strain rate in the southern Indian peninsula. A panel discussion chaired by K. N. Khattri (Dehra Dun) was also organized during this meeting to articulate significant new directions in the studies of the lithosphere.

The panel emphasized the fundamental importance of investigating the geodynamic processes at work in the Indian lithosphere. A better understanding of this basic process is essential for the design of illuminating scientific approaches to a host of important societal problems, notably, quantification of hazards posed by earthquakes, landslides and degrading environments of our resource bases such as ground water aquifers, hill and coastal terranes and biodiversity.

These problems being necessarily complex call for a systematic and synergetic application of investigative tools and approaches from a wide spectrum of earth science disciplines: geological structure and tectonics, geochemistry, geomorphology, geochronology, geothermics, and deep geophysical probes of geodesy, seismology and electromagnetics. In turn, it would be necessary to employ the most incisive statistical and information-theoretic approaches for the analysis and interpretation of a highly diverse yet related data sets.

However, whilst the desirability of investigating the key elements of the entire Indian lithosphere is quite clear, it would be prudent to concentrate the first systematic integrated efforts in defining the structure of its northern boundary – the Himalayan region – and the style and rate of geodynamic processes operating there. For, the grand scale of the tectonic processes in the Himalayas, providing a fast rate of information flow, not only offer an extraordinary opportunity for investigating their nature in a relatively shorter period but also mark this region, with dense population at its southern boundary, as one facing the highest degree of natural hazard. Creation of illuminating knowledge bases to address this problem of great social concern is, therefore, manifestly urgent.

The databases in various disciplines mentioned above are being developed by various National agencies: the IMD, GSI, NGRI, WIHG and university departments of geosciences. The creation of a modern data information system with exchange facilities would greatly facilitate integrated analysis of this data set and promote its wider simultaneous use for greater purpose by diverse groups of scientists in the country. Also, high-quality data sets generated by several international agencies, notably, digital wide-band, wide-dynamic-range seismograms and gravity and magnetic data measured from satellites, could then be progressively added to this system to foster creative approaches to the study of geodynamics of the Indian lithosphere. The panel accordingly recommended creation of a National Geoscience Information System for abstracting, structuring, quality testing and archiving the corpus of data available from national and international agencies, and facilitating its use by Indian scientists to address urgent problems of development and hazard mitigation. Such a system is conceived as being comprised of a number of component data centres in appropriate institutions equipped with quality control and archiving facilities, but integrated with a communication link. It was further recommended that collaborative programmes for developing new conceptual models be formulated, backed by incisive experiment design for generating new data sets. These proposals may then be submitted to DST, UGC, AICTE, etc., for funding. It was also recommended that some centres may be identified and supported for intensified research programmes in the new developing approaches using mathematical modelling and computer simulation.

Some approaches suggested for an effective implementation of these programmes are given below:

- Design of integrated geoscientific investigations to define the structure and tectonics of critical segments of the Indian lithosphere, with immediate focus on its northern collision boundary, using various tomographic approaches, and modelling of geomorphic, magmatic and tectonic processes.
- Development of adequate scientific expertise in the use of wide-band–wide-dynamic-range seismological data and their analysis using waveform modelling techniques.
- Development of nonlinear models of lithospheric processes and their implications to those in the crustal layer.
- Systematic deformation monitoring in critical areas of the Indian crust using GPS, and delineation of areas of anomalous strain accumulation and potential seismicity as well as quantification of strain rates.
- Nucleation of computer-aided Geoscience Information Systems at local scales in various institutions as a prelude to their subsequent integration into a National Geoscience Information System.

Ramesh P. Singh, Indian Institute of Technology, Kanpur.

Necessity of broadening science base in India

This has reference to articles published in *Current Science* about 'ills of Indian Science', 'Indian Science being drifted into obsolescence' and totally 'getting drowned in the sewage', etc. I would like to point out that such articles, in addition to having a demoralizing effect on young Indian scientists, do not reveal the reality. For instance, I feel that my branch of science (organic chemistry) is better now in India than it was 20 years before and the gap between us and the western scientists is getting smaller in this area. This may also be true in several other areas.

These facts are not insignificant when one considers the reality that the candidates admitted to S&T institutions in India come mainly from a small fraction of the 900 and odd million. So, we should compare ourselves only with countries such as Finland, New Zealand, etc.

If we wish to compete with scientists in USA, Japan, etc., then it is important to improve on the competitive level of scientists in India by making the S&T institutions accessible to candidates from the entire population. That is, instead of declining admission in S&T institutions, *de facto*, to a large section (> 95%) of our population, both the political and science policy makers should come out with plans so as to tap talent from the entire population. This would certainly improve the competitive level of Indian scientists. The Chinese have amply demonstrated this in the recent years.

For example, 20 years back, Indian students did not have problems in scoring >90 percentile in GRE examinations. Their competitors were mainly the Taiwanese and south Koreans. But these days their main competitors are mainland Chinese and our toppers from even prestigious institutions get poor 60–70 percentiles! Clearly, the Chinese students who come through a massive competition at home are able to compete better than Indians in the international arena!

MARIAPPAN PERIASAMY

School of Chemistry,
University of Hyderabad,
Central University P.O.,
Hyderabad 500 046, India.

More on the NSU

Apropos the rejoinder of P. N. Srivastava, which appeared in *Current Science*, 1995, **68**, 1190, may I ask the following straight questions to the author?

1. When one wants to build an institution of quality, why should one think of reservation, even for NRIs? Does it not reflect the current Indian mentality, reservation for weaker sections, now reservation for NRIs. Why should the recruitment not be on merit and merit alone?

2. If the Government is to contribute substantially for the establishment of National Science University (NSU), will it not bring pressure for reservation in admissions, and even in staff recruitment, for weaker sections of society?

3. What is the guarantee that NRIs recruited would be meritorious? People who want reservation would also manipulate and make believe that they are good.

4. What is the guarantee that NRIs would continue supporting NSU financially? People who want advantages in the beginning itself would desert the NSU once they find that they are not deriving any benefit from it.

5. Why should the NRIs not help financially the development of science in India, without any strings attached?

6. Why should the NRIs not think of improvements in IISc and in IITs by persuading the Government to change their constitution and pouring money in them to equip them well with staff and equipment? Many of the NRIs are products of these institutions and they will only be paying their debt to their Alma Mater. The NRIs, of course, would not be satisfied, for they would not have a dominant hold on the institution.

7. What is the guarantee that the NSU-trained graduates and postgraduates will have the missionary zeal to work, on relatively ill-paid salary, in colleges and universities which would continue to be ill-equipped as compared to the NSU.

8. Will not the NSU-trained graduates and postgraduates seek greener pasture in USA and other advanced countries?

9. Is the present concern of the NRIs for science in India because (to quote from Mahajan and Sudarshan (see *Current Science*, 1995, **68**, p. 1260 (1295))

(i) he does not feel a full citizen of the institution in America which controls his career,

(ii) he, having gone knocking on the American doors, believing in the universality of science, has now realized, although the realization has come rather late in the day, that American scientific establishment is less than fair to many of its most outstanding imports.

In my opinion the crying need of the hour is not setting up new institutions but strengthening selected existing institutions, such as IISc, some IITs, some universities, massively with funds, additional staff of quality and also introducing some reforms in the administration.

There are many Indians working in India who are intellectually much better equipped than most of the NRI scientists.

N. S. NARASIMHAN

Centre for Agro Chemical Research,
SPIC Science Foundation,
Mount View,
110, Mount Road, Guindy,
Madras 600 032, India.

Response:

The concerns expressed by Narasimhan are a clear case of misinformation or ignorance towards facts about which I had referred to in my rejoinder. There is, therefore, no doubt why some persons have referred to the whole project as 'naive' because they have not read my articles 'The concept of the National Science University: A rejoinder' and 'National Science University: Final Report' (*Current Science*, 1995, **68**, 1190–1198) carefully. I respond to the specific questions pointwise.

1,3. I wonder from where the author has got the idea that there will be reservations for NRIs. There will be no reservations for anybody, NRIs or Indians. The only point that has been made is that 'Further, what is wrong to expect that about 20–25 (out of a faculty strength of 200–250) well-established Indian scientists working abroad should decide to return to India for good to give a boost to this university by working here' (p. 1190). This is no reservation. 'The selection will be done with the active participation of and advice from the International Advisory Committee (p. 1194)' only on merit and no other consideration. Is this reservation?

2. As already mentioned in point number 1, there will be no reservations in admissions, staff recruitments, etc. The point is very clear to us that we would rather not have the university than have a university with reservations. The financial contribution from the government is not sacrosanct. The objectives of the university will, in no case, be sacrificed.

4. The university will be established only when the trust will be formed and there is no need to have any guarantee from the NRIs. The responsibility of raising further funds will be that of the trust.

5. No contribution or donation with strings will be accepted either from Indians or from NRIs.

6. Why should IISc or IITs not approach the NRIs for financial assistance? Who has put this restriction? Who can allege that more money has not been poured in IISc and IITs as compared to the universities? When I was in Bangalore during the last Indian Institute of Science (Bangalore) Court Meeting, a number of scientists belonging to IISc questioned the money that was poured into the Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore, and stressed that if only a part of that money would have been allocated to IISc, it would have improved the Institute immensely. Why did no one raise this issue at that time?

7,8. Why should there be any guarantee for this? Is it not a fact that many graduates of the IISc, IITs, many universities such as Delhi, JNU, BHU, Hyderabad, etc., stay on in India and has the country not benefited from them? Some of them do go out for good but others return as well. If the argument of the author is carried to the other end then only those institutions should be supported whose graduates are considered unfit for selection for going abroad. However, I have always been positive and do hope that Indian colleges and universities are going to improve in future.

9. This point refers to an article published in 1986 and reprinted in *Current Science*. The views expressed in this article have no bearing on the establishment of the National Science University. The pertinent official documents in this connection prepared by the committees appointed by the Government of India are the Concept of the National Science University and the Structure of NSU, reference to which has been made in *Current Science*, 1995, **68**, 1192.

P. N. SRIVASTAVA

Nuclear Science Centre,
P.O. Box 10501,
JNU Campus,
New Delhi 110 067, India.

SCIENTIFIC CORRESPONDENCE

Comments on 'What kind of a system is a human being? A cybernetic system?' [*Curr. Sci.*, 1994, 67, 936–940]

Narasimhan's account of the history of research in human cognition and his prognosis as to its future are well taken. Of concern, however, is his use of the term 'cybernetic system' without definition. Neither Wiener nor, as far as we know, Ashby, ever used this expression. The word 'system' was used by Wiener only in familiar contexts such as the solar system or the nervous system, and he never thought of himself as a 'system theorist'. Indeed, the technical uses of the term 'system' in the literature make it equivocal, and there have been complaints about 'fuzzy thinking' in this field, e.g. from Kalman *et al.*[1]. It is, therefore, essential that the term 'cybernetic system' be clearly demarcated before the question 'Is a human being a cybernetic system?' is asked. Such a demarcation will, of course, depend on how one interprets the underlying concept, *cybernetics*.

Unlike the term 'machine', which Wiener defined clearly as 'a multiple-input, multiple-output transducer', he never clearly prescribed the range of the term 'cybernetics'. Narasimhan [p. 937, col. 1 para 4] mentions Wiener's use of it to mean 'the entire field of communication and control in the animal and the machine'. Since very large tracks of engineering and almost the whole of the human and biological sciences involve communication and control, this definition is woefully general. A fuller rendition of what Wiener intended appears in his 1950 book[2]:

'Since the end of World War II, I have been working on the many ramifications of the *theory of messages*. Besides the electrical engineering theory of the transmission of messages, there is a larger field which includes not only the study of language but the study of messages as a means of controlling machinery and society, the development of computing machines and other such automata, certain reflections upon psychology and the nervous system, and a tentative new theory of scientific method. This larger theory of messages is a probabilistic theory, an intrinsic part of the movement that owes its origin to Willard Gibbs.

Until recently, there was no existing word for this complex of ideas, and in order to embrace the whole field by a single term, I felt constrained to invent one.' [50j, p. 15]

These words suggest that a better approach is to treat cybernetics not as a science *per se*, but rather as a *movement within science*, which draws attention to and focuses on the stochasticity of the cosmos, and on the new concepts that stochasticity allows, such as communication, information, teleological machines, regulation, intelligence, etc. (cf. *Curr. Sci.*, 25 December 1994, pp. 919–920). Calling Leibniz 'the patron saint of cybernetics', Wiener felt that all scientific attempts to absorb mind into nature, such as studies of machine intelligence (chess-playing automata), language, and animal and human intelligence, should come under the rubric 'cybernetical'. The central concepts are 'message' and 'machine' (or monad or black box), i.e. transformer of messages; 'feedback', while important, is not quite that central.

From this perspective, the lines Narasimhan has drawn between the three approaches to human understanding, viz. the cybernetic (in a narrow sense), the artificial-intelligence, and the connectionist, appear to be too sharp. These approaches should be seen as steps in an ongoing effort, that is cybernetical in the broad sense. This effort, the ultimate aim of which must be the understanding of the entire human being, should recognize that this being is 'the most complicated object under the sun', as von Neumann put it. Today this effort is rightly focused on understanding the 'Homo sapiens–faber' part of man. However, his conspicuously active but more difficult 'peccator' part should not be entirely overlooked.

1. Kalman, R. E., Falk, P. L. and Arbib, M. A. , *Topics in Mathematical Systems Theory*, McGraw Hill, New York, NY, 1969.
2. Wiener, N., *The Human Use of Human Beings*, Houghton Mifflin, Boston, MA 1950, Da Capo Series in Science, Plenum, New York, NY, 1988.

P. R. MASANI

Department of Mathematics and Statistics, University of Pittsburgh, Pittsburgh, PA 15260, USA.

Comments on the review of the monograph *The Scientific Methodology in the Light of Cybernetics* [*Curr. Sci.*, 1994, 67, 954–956]

Narasimhan's review contains misunderstandings as to what the term 'scientific methodology' means, and unfounded charges of omission and of exaggeration on my part.

The scientific methodology (SM) is the body of principles governing the method of inquiry used by scientists both to acquire knowledge about the world and to apply it practically. It is intended to provide the practising scientist with enough principles to pursue his craft vigorously, and to offset any doubts he may entertain as to its worthiness. The SM has, first of all, to state what it is in the world that makes science possible, and to describe its classifications and its different aspects, such as experimentation, reasoning, concept formation, mathematicization, learning from the past, and the mode of discovery. Secondly, the SM must say how the notions of message, communication, intelligence, teleology, contest, etc., which until fairly recently were left out of the scientific arena, are scientifically demarcatable. Thirdly, the SM must expose instances of faulty conceptualization. Fourthly, the SM has to say how scientific knowledge is related to the knowledge of craftsmen, philosophers and religionists. Finally, the SM must comment on how science bears on human welfare. Let us from this standpoint judge the blemishes that Narasimhan finds in his review [R] of the monograph [M] in question. They fall under 14 heads, which I have labeled I, II, ... ,XIV.

I. In [R, p. 954, para 2], Narasimhan writes that the following have not been addressed in [M]:

'(i) What is science about? (ii) What is the logical structure of science? (iii) What are the distinguishing characteristics, if any, of the activities that create science? (iv) What is the difference, if any, between scientific knowledge and commonsense knowledge? (v) Are the structures of these two kinds of knowledge similar? (vi) Do these two kinds of knowledge relate to the same world? (vii) What is the relationship between the language of commonsense and the language of science?'

Of these, (vi) is nonsensical since there is only one kind of true knowledge, that of the one world in existence. Of the rest, (i) is a little too vague to be of much significance. However, an adequate answer, to wit, *science is about understanding the world around us by the method of combining logic, mathematics and observation, which came with the Pythagoreans*, is latent in the following passage:

'The purpose of the enterprise throughout has been to understand the world around us. However, what made the work of Newton possible was the methodology of combining mathematics

and observation which came with the Pythagoreans, and the consistency with which the great successors of Pythagoras have practised it.' [M, p. 19]

The question (ii) is, of course, most germane to the SM, but it is already addressed in [M, §3]. The logical structure rests on Einstein's dictum: 'As far as the laws of mathematics refer to reality they are not certain; and as far as they are certain, they do not refer to reality', and can be summed up in the four basic principles:

PI : All the statements of empirical science, including statements of natural laws, are uncertain hypotheses.

PII : (a) The truths of logic and mathematics are devoid of factual content; (b) their certainty stems solely from the rules of language.

PIII : The truths of logic and mathematics provide the deductive vehicle which allows the transition from general hypotheses to empirically testable statements.

PIV : The concepts and hypotheses of science are objective in that each has a content exceeding that of observational terms or statements. But these concepts and hypotheses are not unique, and the choice of one set in preference to another is a human imposition. The imposition has been found to be inconsequential: the resulting theories most often turn out to be equipollent. Finally, the judgement of the significance of the observational errors is made by human convention.

Narasimhan does not say where these principles fall short, or what has to be added to complete them, that is not already in [M].

The question (vii), again germane to the SM, is addressed in [M, §4]. The answer in brief is: *the useful concepts of everyday language (e.g. blue, hot, angry) are refined and elevated into the scientific vocabulary by a process called explication* (Carnap's term). The question (iii) is somewhat vague. Obviously, eating or sleeping are not scientific activities (although practised in moderation, they are needed to 'create science'), and, obviously, observing, experimenting, making, concept forming, deducing, and abducting are. The question (iv) has the easy answer: *scientific knowledge is a refinement of commonsense knowledge*. As for question (v) the term 'structure' is not applicable to 'commonsense knowledge', until we know what the latter means exactly. Hence, as it stands, (v) like (vi) is meaningless.

To sum up, Narasimhan is wrong in saying that the two questions (ii) and (vii), which are germane to the SM, are not addressed in [M], and is himself at fault for formulating two, viz. (iii) and (iv), which are trivial and easily answered by what is in [M], and two more, (v) and (vi), whose meaningfulness is in doubt.

II. In [R, p. 954, para 3] Narasimhan states that I do not discuss in a 'systematic manner' the relationship between *pre-scientific* and *post-scientific technologies*. But what is germane to the SM is not this relationship but the one between craft (i.e. pre-scientific technology) and science, and on this enough is said in [M]. First, the role of the craftsman is emphasized in a paragraph with the heading *The Craft as Starting Point*, wherein is emphasized the input of the craft in leading to sound concept formation and thereby to the initiation of new sciences, for 'the craftsman intelligently cooperates with nature' [M, p. 35]. Such input continues to this day [M, p. 51]. Secondly, the craftsman is invaluable to science in providing good apparatus to the experimenter. Indeed, 'for a science to emerge, either Nature has to be generous in providing enough clues to the inquirer, or else a craftsman has to give him what he may need in order to experiment intelligently' [M, p. 20]. Briefly, the SM 'rests on man being both *Homo sapiens* and *Homo faber*' and '... without the craftsman, the scientist is lost' [M, pp. 50, 51]. As to the differences between the earlier (pre-scientific) technologies and science-based technologies of the last 200 years, the answer obviously is that the latter are on the whole much superior, especially in the military field. Consequently, societies that are unable to shake off the former and shift to the latter get dominated by those that do – a fact important for human history but not directly for the SM.

III. In [R, p. 954, col. 1, para 4] Narasimhan writes that the issues discussed in [M] (summed up in para 2 of this letter) have only 'a tangential bearing on ... the SM', and in [R, p. 954, col. 2, para 2] he suggests that I have tried 'to project him [Wiener] as a man for all time'. But until Narasimhan can substantiate these charges by citation of page and para from [M], neither claim nor suggestion can be taken seriously. His remaining comments in these paragraphs as to who influenced me and what I did for 40 years are irrelevant to the SM.

IV. In [R, p. 954, col. 2, para 3] Narasimhan speaks of my problematic 'attitude to cybernetics as a framework for analysing purposeful behaviour in animals and humans' but without saying what this attitude is. In [M] the noun 'cybernetics' hardly ever appears after the Preface, although the adjective 'cybernetical' does. In the Preface [M, p. 8] the 'cybernetical perspective' is described as one in which the concepts of *message, communication* and *control* are paramount. True, that on [M, p. 82], I give the Bigelow–Rosenblueth–Wiener definition of a *teleological mechanism*. But on the same page I also point out that there is an 'important difference between engineered teleological mechanisms and living systems'. Thus, this 'attitude' is some ghost of Narasimhan's own making, which needs exorcising*.

V. In the same paragraph Narasimhan goes on to say that I overstretch the 'scope of cybernetics' and 'trivialize its technical terminology' by stating that in social undertakings 'moral evil should be construed as teleological noise'. But what is incorrect about this statement? The following illustration taken from a recent lecture makes its truth crystal-clear:

'Consider a department store as an input–output transducer, the output terminals of which are sales desks operated by well-paid clerks. The clerk weighs a pound of candy for a customer, but pilfers a few pieces while packaging under the counter. Obviously, such pilfering has a dissipative effect on the successful operations of the store, and thus plays a role analogous to that played by natural noise in, say, an electronic transducer. Unlike natural noise, however, this new 'noise' of human origin, is teleological. It is created in order to fulfil a purpose, viz. to pilfer, and this desire to pilfer is a simple example of the moral evil. The filtration of this noise, unlike that of natural noise, will involve a contest between the sales clerks and the store owner and his detectives. We are led to the following:

*In the 25 December 1994 issue, Narasimhan has written an interesting paper, but it is marred by his use of the term 'cybernetic system' without saying what it means (see the previous letter).

Triviality: (a) In social undertakings the dissipation or noise caused by moral evil is teleological; (b) the filtration of natural noise is contest-free; the filtration of teleological noise requires contest.'

These issues are clearly latent in Wiener's writings on noise and evil.

VI. In the same paragraph, Narasimhan writes that I make 'exaggerated claims' in calling the study of the genetic code, molecular biology and genetic engineering 'cybernetical territory par excellence' [M, p. 83]. But again he is wrong, for these subjects are heavily involved with communication (information), control, and (other Wiener favourites such as) teleology, Maxwell's demons, self-organizing and reproducing mechanisms, as a glance at the Contents of Monod's *Chance and Necessity* (1970), Chapters I, III, IV, shows:

I Of strange objects: The natural and the artificial ... Difficulties of a space programme ... Objects endowed with a purpose ... Self-constructing machines ... Self-reproducing machines ... Strange properties: invariance and teleonomy ... The 'paradox' of invariance – Teleonomy and the principle of objectivity.

II Maxwell's demons: Proteins as molecular agents of structural and functional teleonomy ... The enzyme proteins as specific catalysts ... Covalent and noncovalent bonds ... The concept of the noncovalent stereo-specific complex ... Maxwell's demon.

IV Microscopic cybernetics: Functional coherence of cellular machinery ... Regulatory proteins and the logic of regulations ... Mechanism of allosteric interactions ... Regulation of the synthesis of enzymes ... The concept of gratuity ... 'Holism' vs. 'reductionism'.

VII. In [R, p. 954, col. 2, last para] Narasimhan questions my assertion that with regard to classification 'the scientific tradition is close to Aristotle'. But indeed it is, for Aristotle allowed 'the natural contours within the cosmos, as perceived by the mind ... to demarcate the sciences' [M, p. 95]. Obviously, since the contours perceived by a scientific mind today will be different from those perceived by even the greatest mind in 400 BC, the demarcations must change. Furthermore, a perusal of Joachim's Introduction to the *Nichomachean Ethics* will show that Aristotle saw the need for interdisciplinary work, Hence, this issue is spurious.

VIII. The statement 'science is based on faith', which heads [R, p. 954, col. 3] is a corollary of two more basic principles of the SM:

1. Science rests on the assumptions that the world is orderly and that the human mind can comprehend this order.
2. These assumptions are based on extralogical faith; cf. [M, §1]

The faith in (2) is none other than the *primal faith* on which all life depends (cf. [M, pp. 13–14, p. 16, last para]) and the religionist's faith in God '*presupposes an orderly and discoursable world and so depends on a primal faith that such is the world we live in*' [M, p. 17]. Thus, Narasimhan is wrong in saying that I claim that the primal faith is 'of the same kind as the religious faith of medieval schoolmen' [R, p. 954, last para]. The truth [M, p. 17] is that the primal faith is *more basic than* the 'religious faith' of medieval schoolmen, Vedantists and many others, including scientists; cf. [*Curr. Sci.*, 25 December 1994, pp. 917–919].

IX. Wiener has written: 'Without faith that nature is subject to law there can be no science. No amount of demonstration will prove that nature is subject to law'. But since Narasimhan does not tell us the flaw in Wiener's full argument and that of A. N. Whitehead [M, pp. 13, 14], his skepticism as to whether 'the extralogical faith of a scientist is a faith' [R, p. 954, last para] remains to be justified. Notwithstanding his skepticism of this faith, Narasimhan attaches great importance to 'the *origin* of it' [R, p. 954, col. 1, para 2]. But once the orderliness of the world is accepted by faith, its origination, while important to the child psychologist, is of zero import to the SM, for life itself, and *ipso facto* a discussion in child psychology or in any other field, would be impossible without such orderliness. In short, *one can investigate the origin of the primal faith only by practising, it.* Hence, Narasimhan's three paragraphs about Piaget are irrelevant to the SM*.

X. Regarding the notion of 'scientific revolutions' and of 'paradigm shift' (cf. [M, pp. 20–21]) Narasimhan misleads by saying 'Masani argues' [R, p. 954, col. 3, middle para] instead of saying 'Masani quotes Heisenberg, Einstein and Schrödinger to show that'. It was Heisenberg who said that the 'revolution' (*ca.* 1600) was '*a transition from the descriptive science of Aristotle to the structural science of Plato*' [M, p. 23], and Einstein who said that the relativity 'revolution' (*ca.* 1915) was '*no revolutionary act but the natural continuation of a line that can be traced through centuries*'; and on the quantum 'revolution' (*ca.* 1925) it was Schrödinger who described the wave mechanics as '*a thorough organic expansion and development, one might almost say merely a restatement of the old* [Hamilton, Jacobi] *theory in more subtle terms*' [M, p. 22]. An examination of how these great scientists do their research would show that they strive to complete and beautify the earlier mathematical equations in the light of the new experimental data, and that paradigms play no role whatever in their creative work. It is this that is relevant to the SM.

XI. Narasimhan [R, p. 955, col. 2, paras 1, 2, 4, 5] addresses the fundamental question of *abduction*, i.e. the process by which the mind is able to avoid the huge number of hypotheses that merely fit the known facts, and to hit upon concepts and propositions which yield new knowledge, i.e. *laws of nature*. All except paragraph 5 tell the true story. Paragraph 5, however, contains the incorrect statement that 'the theoretical underpinnings' (more accurately, scientific explications) of 'terms like instinct, insight, abductive skill are glossed over' in [M]. The need for such explications, far from being glossed over, has been clearly emphasized; thus:

'It remains to be discovered what factors in the human nervous system account for man's prodigious criticality in relation to learning automata, and whether Wiener's idea of grading automata in a hierarchy of types offers a fruitful approach to such study. More generally, *a scientific study of what Peirce called intunement should be on the cybernetical agenda.*' [M, p. 84]

'The far greater flexibility and criticality of biological organisms *vis-à-vis*

*In the author's earlier paper 'The illusion that man constructs reality' (*Kybernetes*, 1992, **21**(4), 11–24), the obvious contributions of Piaget to child psychology are lauded, and the dubiety of their relevance to the SM laid bare. Without the sun and the earth there would be no babies; thus, cosmic order first, babies second. Babies do not 'construct' the world; they slowly discover its order.

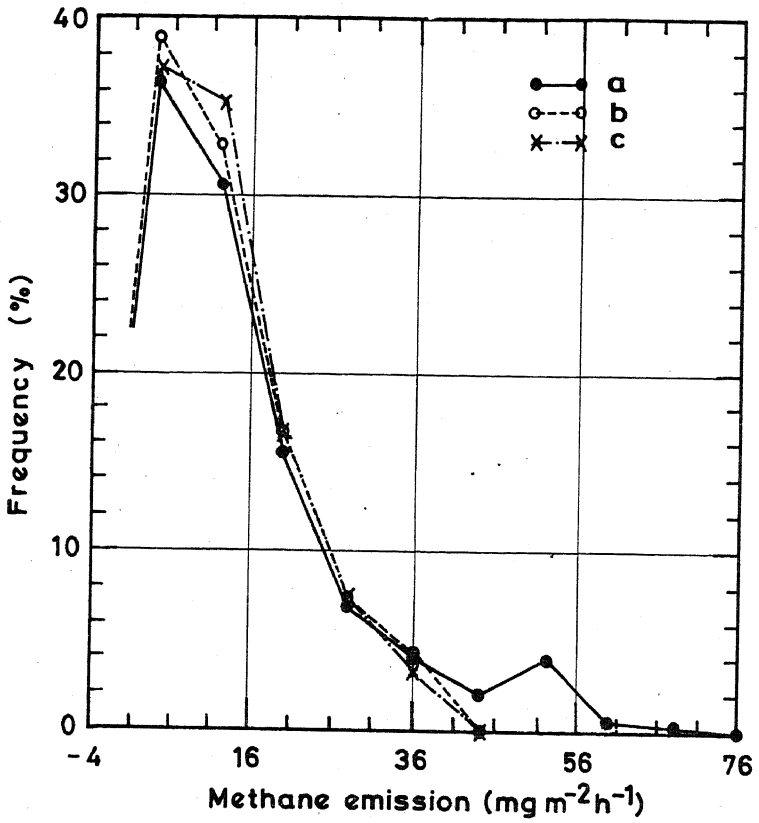


Figure 1. Frequency polygon for CH_4 emission rates. Each data point is an average of several but variable number of replicates. ***a***, all data, ***b***, exceptionally high values due to high inputs of chemical fertilizer/organic matter excluded; ***c***, values from all pot experiments excluded in addition to exceptionally high values from field experiments due to high inputs of fertilizer/organic matter.

Table 1. Calculated CH_4 emission from rice paddies on the basis of data in Figure 1

	Emission rate (mg CH_4 m^{-2} h^{-1})	Annual emission (Tg CH_4 yr^{-1})
All data		
Simple average	14.89 (±12.95)	49.13
Geometric mean	9.32	30.76
Exceptionally high values removed[1]		
Simple average	12.29 (± 8.88)	40.58
Geometric mean	8.24	27.19
Values for all pot experiments removed[2]		
Simple average	12.34 (± 8.14)	40.72
Geometric mean	8.97	29.60

Values in parentheses are standard deviations.
[1]Values from experiments with heavy doses of chemical fertilizers and paddy straw/horse manure were excluded. Most of these were pot experiments.
[2]Values from all pot experiments involving chemical fertilization and organic matter amendment were excluded.

around the world[6] and averages 105 days. Using 1.31×10^{12} m^2 rice cultivation area, a CH_4 emission period of 105 days and average emission rates, the total global emission varies from 40.58 to 49.13 Tg yr^{-1} (Table 1). Because of the pronounced skewness in the data and high standard deviations associated with simple averages, it would seem preferable to use geometric means. Based on geometric means, the total CH_4 emission varies from 27.19 to 30.76 Tg yr^{-1} (Table 1).

Inclusion of the data from artificial pot experiments or from field experiments with very high inputs of organic matter in calculations of global emission may not be realistic. A recent review indicated that the CH_4 flux response to the application of fertilizer or/and organic matter ranged from negligible to dramatic (–75 to +6857% compared to control), and varied from year to year, among periods within a growing season, and from treatment to treatment[38]. Additionally, if harvested wetland rice area of 1.236×10^{12} m^2 is used as calculated by Neue and Roger[37], the global estimate reduces to 25.68–27.84 Tg CH_4 yr^{-1}. We suggest that these values could be considered a reasonable estimate until all rice-growing regions and rice ecologies are adequately covered by flux measurements. These estimates depend much on the CH_4 emission period, which may vary substantially in different geographic regions. Assuming a 30 d emission period, as argued by Sinha[39], the estimate for global emission reduces to 7.34–7.95 Tg yr^{-1}. This estimate compares with 7.08 Tg yr^{-1} assessed by Sinha[39] on the basis of rice yield and 30 d CH_4 emission period. Thus, there is a need for substantial reduction in the global estimate of CH_4 emission from rice paddies.

1. Crutzen, P. J., *Environ. Monitor. Assess.*, 1994, **31**, 1–15.
2. Watson, R. T., Meira Filho, L. G., Sanhueza, E. and Janetos, A., in *Climate Change 1992*, the Supplementary Report to the IPCC Scientific Assessment (eds Houghton, J. T., Callander, B. A. and Varney, S. K.), Cambridge Univ. Press, Cambridge, 1992, pp. 25–46.
3. Aselmann, I. and Crutzen, P. J., *J. Atmos. Chem.*, 1989, **8**, 307–358.
4. Khalil, M. A. K. and Shearer, M. J., in *Atmospheric Methane: Sources, Sinks and Role in Global Change* (ed. Khalil, M. A. K.), Springer, Berlin, 1993, pp. 180–198.
5. Matthews, E., Fung, J. and Lerner, J., *Global Biogeochem. Cycles*, 1991, **5**, 3–24.
6. Shearer, M. J. and Khalil, M. A. K., in *Atmospheric Methane: Sources, Sinks and Role in Global Change* (ed. Khalil, M. A. K.), Springer, Berlin, 1993, pp. 230–253.

7. Schütz, H., Holzapfel-Pschorn, A., Conrad, R., Rennenberg, H. and Seiler, W., *J. Geophys. Res.*, 1989, **94**, 405–416.
8. Holzapfel-Pschorn, A. and Seiler, W., *J. Geophys. Res.*, 1986, **91**, 11803–11814.
9. Watanabe, A., Murase, J., Katoh, K. and Kimura, M., *Soil Sci. Plant Nutr.*, 1994, **40**, 221–230.
10. Watanabe, A., Katoh, K. and Kimura, M., *Soil Sci. Plant Nutr.*, 1993, **39**, 707–712.
11. Watanabe, A., Katoh, K. and Kimura, M., *Soil Sci. Plant Nutr.*, 1994, **40**, 497–504.
12. Watanabe, A., Katoh, K. and Kimura, M., *Soil Sci. Plant Nutr.*, 1993, **39**, 701–706.
13. Parashar, D. C., Rai, J., Gupta, P. K. and Singh, N., *Indian J. Radio Space Phys.*, 1991, **20**, 12–17.
14. Jermsawatdipong, P., Murase, J., Prabuddham, P., Hasathon, Y., Khomthong, N., Naklang, K., Watanabe, A., Haraguchi, H. and Kimura, M., *Soil Sci. Plant Nutr.*, 1994, **40**, 63–71.
15. Nugroho, S. G., Lumbanraja, J., Suprapto, H., Sunyoto, Ardjasa, W. S., Haraguchi, H. and Kimura, M., *Soil Sci. Plant Nutr.*, 1994, **40**, 275–281.
16. Saha, A. K., Rai, J., Raman, V., Sharma, R. C. and Parashar, D. C., *J. Radio Space Phys.*, 1989, **18**, 215–217.
17. Gupta, M., Verma, S. D., Parashar, D. C. and Gupta, P. K., *J. Radio Space Phys.*, 1994, **23**, 265–268.
18. Yagi, K., Chairoj, P., Tsuruta, H., Cholitkul, W. and Minami, K., *Soil Sci. Plant Nutr.*, 1994, **40**, 29–37.
19. Cicerone, R. J., Shetter, J. D. and Delwiche, C. C., *J. Geophys. Res.*, 1983, **88**, 22–24.
20. Cicerone, R. J., Delwiche, C. C., Tyler, S. C. and Zimmerman, P. R., *Global Biogeochem. Cycles*, 1992, **6**, 233–248.
21. Seiler, W., Holzapfel-Pschorn, A., Conrad, R. and Scharffe, D., *J. Atmos. Chem.*, 1984, **1**, 241–268.
22. Yagi, K. and Minami, K., *Soil Sci. Plant Nutr.*, 1990, **36**, 599–610.
23. Dai, Aiguo, Methane emission rates from rice paddies in Hangzhou China during fall of 1987, MS Thesis, 1988.
24. Schütz, H., Seiler, W. and Rennenberg, H., in *Soil and the Greenhouse Effect* (ed. Bouwman, A. F.), Wiley, New York, 1990, pp. 269–301.
25. Khalil, M. A. K. and Rasmussen, R. A., in *Proceedings of the Indo-US Workshop on Impact of Global Climate Changes on Photosynthesis and Plant Productivity* (eds Abrol, Y. P., *et al.*), Oxford and IBH Publishing Co., New Delhi, India, 1991, pp. 641–650.
26. Sass, R. L., Fisher, F. M., Harcombe, P. A. and Turner, F. T., *Global Biogeochem. Cycles*, 1990, **4**, 47–68.
27. Sass, R. L., Fisher, F. M. and Harcombe, P. A., *Global Biogeochem Cycles*, 1991, **5**, 275–287.
28. Sass, R. L., Fisher, F. M., Turner, F. T. and Jund, M. F., *Global Biogeochem. Cycles*, 1991, **5**, 335–350.
29. Sass, R. L., Fisher, F. M., Wang, Y. B., Turner, F. T. and Jund, M. F., *Global Biogeochem. Cycles*, 1991, **6**, 249–262.
30. Chen, Z., Debo, L., Shao, K. and Wang, B., *Chemosphere*, 1993, **26**, 239–246.
31. Cicerone, R. J. and Shetter, J. D., *J. Geophys. Res.*, 1981, **86**, 203-207, 209.
32. Khalil, M. A. K. and Rasmussen, R. A., *Tellus*, 1989, **B42**, 229–236.
33. Wassman, R., Papen, H. and Rennenberg, H., *Chemosphere*, 1993, **26**, 201–218.
34. Mariko, S., Harazono, Y., Owa, N. and Nouchi, I., *Environ. Expt. Bot.*, 1991, **31**, 343–350.
35. Parashar, D. C., *et al.*, *Curr. Sci.*, 1994, **66**, 938–940.
36. Kreileman, G. J. J. and Bouwman, A. F., *Air Soil Poll.*, 1994, **76**, 231–258.
37. Neue, H.-U. and Roger, P. A., in *Atmospheric Methane: Sources, Sinks and Role in Global Change* (ed. Khalil, M. A. K.), Springer, Berlin, 1993, pp. 254–298.
38. Singh, J. S. and Singh, Smita, *Trop. Ecol.*, 1995, **36** (in press).
39. Sinha, S. K., *Curr. Sci.*, 1995, **68**, 643–646.

ACKNOWLEDGEMENT. This study was funded by the Ministry of Environment and Forests, Government of India, New Delhi.

J. S. SINGH
SMITA SINGH

Department of Botany,
Banaras Hindu University,
Varanasi 221 005, India.

Nomenclature of super heavy elements

Scientists of different countries have made different claims on synthesis of new elements. As their methods of identification were different, the criterion for discovery of elements[1] remained under dispute for a long time. Hence, the literature has reports on the same element with different names and symbols, which leads to confusion. Recent changes by IUPAC[2] in the names of elements 101–109 have evoked mixed reactions, especially in USA. This is because some old names have been reassigned different atomic numbers, in an effort to solve the dispute.

Some actinide elements were named after the place of discovery in accordance with their rare earth counterparts. For example, eka-Europium was named Americium, and eka-Terbium was named Berkelium. It is interesting to note that eka-Hafnium (104) has aptly been christened after Dubna, just as Hafnium was named after Copenhagen (Hafinia in Latin). The elements with atomic numbers 105 and above have been named after eminent scientists. (105 – Joliotium (Jl); 106 – Rutherfordium (rf); 107 – Bohrium (Bh); 108 – Hafnium (Hn); 109 – Meitnerium (Mt)). The earlier name Seaborgium suggested by the Berkley group for element 106 has not been accepted by IUPAC because it does not favour naming the elements after scientists who are alive. Some subtle changes have occurred with respect to symbols: Mv, Lw and Ha have now become Md, Lr and Hn, respectively[3–5].

Elements yet to be ratified are given temporary appellation derived from the following unique Greek/Latin roots[6]: 0 = nil; 1 = un; 2 = bi; 3 = tri; 4 = quad; 5 = pent; 6 = hex; 7 = sept; 8 = oct; 9 = enn. Instead of this system, a simpler representation with Z is also followed. For example, $(110)O_2$ is the dioxide of the element ununnilium (Unn).

Worldwide acceptance of the recent IUPAC recommendations[7] would end the confusion that prevails in the naming of these 'superheavy' elements.

1. Harvey, B. G., *et al.*, *Science*, 1976, **193**, 1271–1272.
2. IUPAC Announcement, *Indian J. Chem.*, 1995, **A34**, 164.
3. Ghiorso, A., *et al.*, *Phys. Rev.*, 1955, **98**, 1519.
4. Ghiorso, A., *et al.*, *Phys. Rev. Lett.*, 1961, **6**, 475.
5. Ghiorso, A., *et al.*, *Phys. Rev. Lett.*, 1970, **24**, 1503.
6. Fernelius, W. C., *et al.*, *J. Chem. Educ.*, 1975, **52**, 583.
7. IUPAC Nomenclature of Inorganic Chemistry 1.1, in *CRC Handbook of Chemistry and Physics*, CRC Press, Florida, 1994, 74th edn.

S. K. ARYA
S. S. VASAN

Chemistry Group,
B.I.T.S.
Pilani 333 031, India.

Whose life is it anyway? A review of some ideas on the issue of animal rights

Anindya Sinha

National Centre for Biological Sciences, TIFR Centre, P. O. Box 1234, Indian Institute of Science Campus, Bangalore 560 012, India

Sing, dilly dilly duckling, come and be killed,
For you must be stuffed, and my customers filled.

– Old English nursery rhyme

'. . . *violent muscular spasms, occasionally sufficient to throw the animal bodily across the cage occurred, and gradually passed into a state of general weakness ending in death.*' Some animals '*bit themselves severely, two chewing off the end of a finger, and one, the whole skin of the forearm, exposing the muscles from the elbow to the wrist.*'

– From a scientific paper investigating the effect of certain substances injected into the brain of monkeys; *The Lancet*, 19 September 1931.

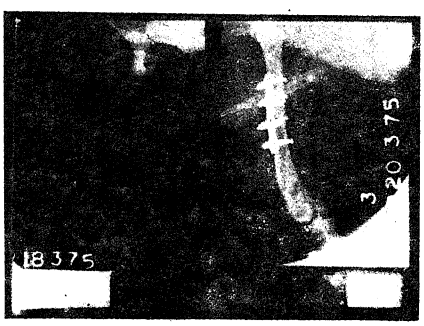
In today's man-eat-man world, a consideration of the exploitation of animals, their rights, and our obligations towards them might seem trivial and fairly unnecessary. Its importance, however, lies in two kinds of arguments. The first encompasses the concept of global ecology and the unique position held by each species on this ecosphere, the Earth. The environmental movement has particularly made us aware of the often irreparable damage that we have inflicted on other species, and made many of us reflect on our duty towards the other inhabitants that share our home. Although extremely important in its own right, I will not concern myself with it in this paper.

The exploitation of animals

The second group of arguments address the questions whether our eating of other animals, our experiments on them and our destruction of their habitats are ethically defensible. This assumes particular importance if one considers the qualitative and quantitative extent to which animals are exploited to carry the burden of our modern-day civilization. The use of animal pelts and skins in articles such as shoes, belts, wallets and watchbands, and in sports goods such as footballs and boxing gloves, the use of animal excreta in fertilizers, the use of animal urine in the manufacture of perfumes and body lotions, and of animal fats in the manufacture of soap, lipstick and chewing gum are so routine and so much a part of our everyday life that few people are concerned or think about it at all. Lack of knowledge also characterizes our most intimate contact with animals, namely, our eating of them. But the fact is that farm animals, in ever-increasing numbers, are being raised in incredibly crowded, unnatural environments according to what are called 'intensive rearing methods'. Animals kept under such conditions lead lives characterized by extreme deprivation, pain and frustration. But raising livestock in this way is the only economically feasible route for the producer to meet the ever-increasing public demand for meat and dairy products. The point I am trying to make is that it is not the producers alone who are responsible for forcing animals to lead the lives they do, we, consumers, too must take a large share of the responsibility.

Even more relevant to most of us in the scientific community is the responsibility we must accept for another





Animals are often subjected to unnecessary pain during scientific experimentation. This dog had one of its limbs broken without any anaesthetic being applied in the course of a veterinary study investigating bone fracture and healing.

major cause of animal expoitation: the use of animals in scientific research. Most laboratory animals are forced to lead lives that are alien to their natures, and moreover, their confinement is characterized by deprivation and pain. This problem is again magnified as increasing numbers of animals are being confined to laboratories around the world. The following statistics should drive this point home. In the United States of America, a study conducted by Rutgers University's College of Agriculture and Environmental Sciences revealed that in the year 1971 alone, American researchers used 15–20 million frogs, 190,000 turtles, 51,000 lizards, 61,200 snakes, 1.7 million birds, 45 million rodents, 700,000 rabbits, 23,000 sheep, 46,300 swine, 200,000 cats, 500,000 dogs and about 85,300 primates[1]. The use of animals as subjects in research seems as commonplace in the scientific community as the use of animal flesh for food in the world at large.

But can such use of animals for human 'benefit' be morally justified? Concern over the treatment accorded to animals is not an abstract theoretical matter. How we treat animals affects us as humans; the quality of animal life affects the quality of human life. Thus, the need to understand and to evaluate various philosophical positions with regard to animal rights and human obligations becomes a matter of practical significance.

Human obligations

Whether man has obligations towards other animals is intimately connected with our belief and knowledge about the capacities that nonhuman animals possess. This stand becomes clear when humans believe that they have certain duties towards other human beings but not towards inanimate objects such as pebbles or clouds because they lack certain properties which human beings possess. Thus, in order to answer the question whether we have duties towards other animals, we have to first answer another question: What capacities must a being have if we are to feel any responsibility towards it?

There are three things some or all of which have been recognized by various thinkers to characterize beings towards whom we, humans, have some duties[2]. These are: (1) rationality, or the capacity to reason; (2) autonomy, or the capacity to make free choices; and (3) sentience, or the capacity to feel pleasure and pain.

The idea that a being must be rational in order that we may have some duties to it actually goes back to Aristotle, though it was very strongly advocated in the thirteenth century by the Catholic philosopher–theologian St. Thomas Aquinas[3]. Aquinas regarded rationality or intellect as a capacity that makes beings more or less perfect. Man, who has some degree of rationality, was thus considered more perfect than other animals, who lacked this capacity. He also believed that so far as our world was concerned, the less perfect beings could be subordinated by the more perfect ones; man thus need have no qualms about eating the lower animals for it is not contrary to nature, as God designed it to be. With regard to animal sentience, Aquinas believed that it was possible to treat animals cruelly since they do have the capacity to feel pain, that is, to cause them unnecessary pain. Although he did think that it was wrong to treat them in this way, he did not believe, for the reason mentioned earlier, that man had any obligation towards them to abstain from treating them cruelly. However, we do have duties towards lower animals insofar as our treatment of them leads us to sin against some rational being, either human beings or God. In fact, Aquinas did believe that people who treat animals cruelly are naturally inclined to treat rational beings in a similar way. We thus do have indirect duties towards other animals, only those borne out of charity.

Aquinas' views can and have been challenged by a number of later thinkers on three principal grounds:

1. Is there any evidence that cruelty to animals leads to cruelty to human beings? For if there is no such evidence, then, according to Aquinas' belief, it would not be wrong to treat animals cruelly; this, in itself, can be considered a debatable point.
2. We could agree with Aquinas that a being must be rational in order for us to have duties towards it, and then point out, as had been done much earlier by the Greek moralist Plutarch[4], that some of the lower animals may have the capacity to reason, although not to the extent that man has. This is a point that has now come to the fore with much of our research in human cognition being focused on comparative studies on the nonhuman primates and other highly social mammalian species.
3. A more fundamental objection disputes the very basis of Aquinas' proposition and questions whether the property of sentience should not be considered a more concrete ground to determine our moral obligations rather than that of rationality. According to this view, again put forward very strongly by Plutarch, and later supported by the English philosopher Bentham[5], it is enough that a being can experience pain and pleasure for us to have a duty towards it of not causing any unnecessary pain. And this is a direct duty, an obligation quite independent of whether our treatment of animals reflects our treatment of other human beings.

Does this mean that it is wrong to kill an animal as a source of food or to inflict pain upon it in the course of a scientific experiment? The answer, according to most thinkers, would appear to be no. Plutarch, for example, concedes that it would not be wrong to kill and eat an animal if doing so is necessary to save a human life. However, he does maintain, and this should strike a familiar chord in a number of us that it is not necessary to eat the flesh of animals to survive or attain sound health. Moreover, he believes that raising and slaughter-

ing of animals are dictated not so much by consideration of health as by that of taste. And since these are obviously unnecessary pleasures for men to indulge in, he goes on to argue that the pain caused to animals by such a treatment is also unnecessary. Thus, although there may be special circumstances in which the consumption of animals is not wrong, Plutarch believes that it is usually so.

Bentham, adopting yet another viewpoint, argues that our treatment of animals would be wrong if the animals suffered more pain than the amount of pleasure human beings received. He, in fact, incorporated the essential basis of moral equality in his utilitarian system of ethics as 'each to count for one and none for more than one'. Although later thinkers have also agreed that everyone's interests should be given such equal consideration, they have not been able to agree on how conflicting interests can be best equated[2]. Another important position in this regard is that held by the renowned missionary and philosopher Albert Schweitzer, who makes it clear that he is, in principle, not opposed to using animals in scientific research, though he speaks very forcefully against their use in 'unnecessary' research[6].

A final viewpoint which cannot be neglected in this discussion is that of Singer[7]. Even at the risk of being labelled an extreme radical, Singer defines most of the above-described ideas as products of our 'speciesist' way of thinking. A speciesist, according to him, is a person who 'allows the interests of his species to override the greater interests of members of other species'. Speciesism thus shares a number of features with sexism and racism, the most important one being a systematic refusal to apply the principles of justice in an equitable fashion. Singer raises the problem of scientific experimentation on animals and argues that the question of whether one would be prepared to let thousands of people die if they could be saved by experimenting on a single animal is akin to asking whether the same experiment could be performed on a human infant. Singer concludes that if the experimenter is not ready to do so, then his readiness to use nonhuman animals is simple discrimination, since adult mice, dogs and monkeys appear to be more aware, more self-directing, and, as far as we can tell, as sensitive to pain as any human infant. According to Singer, there appears to be no relevant characteristics that human infants have but adult mammals do not have to the same or to a higher degree. Moreover, in case of a counterargument that the human infant, if left alone, would develop into more than a nonhuman, he argues that this would still give us no reason for selecting a nonhuman animal rather than a human being with severe and irreversible brain damage as the subject for our experiments. It is thus Singer's belief that the routine use of animals as food and their use in scientific experiments is an expression of unreasonable prejudice and, like racial bigotry, should be fought against with all our compassion and respect for life.

Animal rights

The question of whether animals have rights is intimately related to our understanding of human and animal nature and also to the question of whether we have obligations towards other animals. A major school of thought is of the notion that whenever one being has an obligation towards another, the other being has a corresponding right against the first and *vice versa*[2]. This is known as the correlativity thesis. If we do have some duties towards other animals, it, therefore, necessarily follows that animals have corresponding rights that we must respect.

An interesting criticism of this viewpoint argues differently. It might generally be favourably considered, for example, that we do have certain duties towards preserving our natural ecosystems, or even the great works of human creativity such as the frescoes of the Ajanta caves. Very few would, however, agree that our oceans and forests, as also the Ajanta paintings, have rights. The correlativity thesis may not thus be too widely applicable.

Counteracting the above criticism, proponents of the thesis postulate that it is only the cases of direct duty that should be considered, while the criticism clearly involves examples of duties that are indirect. Thus, if we do have a duty towards our great works of art, it is not something that we owe to them. It is something that we owe to our future generations, for the loss of the Ajanta frescoes would lead to a loss in their enjoyment of beauty. Similarly, our indirect duty towards our woodlands and oceans only symbolizes our direct duty towards our descendants; the correlativity thesis is, therefore, upheld. Feinberg[8] is a notable proponent of this thesis, and has strongly advocated the rights of animals in accordance with our direct obligations to them.

But what is a right? Feinberg believes that to have a right is to have a claim to something or against others. Therefore, if an animal has a right to life, this would imply that it has a claim against others not to take its life. McCloskey[9], on the other hand, defines a right as an entitlement. Thus, if an animal has the right to eat whatsoever it pleases, it is entitled to such food. The difference between the two is that to have an entitlement does not necessarily require one to have any claims against others, while to have a right can be a basis for making such claims. Which of these two analyses is correct? Which one would be more appropriate for nonhuman animals? These are controversial points and unfortunately, such discussions have tended to obscure the main question of whether animals have rights at all, leading the noted English humanitarian Salt[10] to comment, in an article as early as 1912, that 'the controversy over "rights" (is) little else than an academic battle over words'. The situation remains much the same, even today.

Another problem that is often raised is that even if animals do have rights, they lack a language to articulate their claims or entitlements. This, of course, should not pose a major problem since it can always be argued that amongst human beings, very young children, the aged and those mentally deficient are also incapable of articulating their rights; yet we never do dispute their rights.

Yet another question that is important in this regard is what kinds of rights are there at all. At least two kinds of rights can be distinguished – (1) Special rights: these can be acquired because of what we or someone else has done, and not because of what we are. An example is that of a beneficiary to an inheritance. (2) Natural rights: these are what a being has by virtue of being what it is. The rights to life, liberty and happiness are included under this category because they cannot be acquired because of what one does. These are, therefore, inherent rights.

Salt[10] has suggested that animals do have special rights since they are entitled to certain legal rights; a different view is that of the British political philosopher Ritchie[11] who has argued that works of art are also protected by laws, but cannot be conceived to have legal rights of any kind. A more controversial issue, however, is that of whether animals have natural rights. This is not surprising since natural rights are those which a being has by virtue of its nature, and the nature of nonhuman animals is itself very controversial. Hence, the question which has to be raised now concerns the nature of beings that can have rights and whether animals can be considered to be such beings.

Possibly the most prevalent view in this debate is that a being must be rational if it is to be capable of having rights. However, as discussed earlier, a number of thinkers, Aristotle and Aquinas being prominent among them, are of the opinion that animals are apparently incapable of reasoning and, hence, have no rights against us. A major opponent of this argument is Feinberg[8] who has again pointed out that children and the mentally ill may not demonstrate any ability to reason but that does not deprive them of their rights. He has used an argument similar to Singer's in pointing out that even if children are considered to be potentially capable of reasoning, the mentally feeble are definitely not. Hence, if rationality is considered to be a criterion for having rights, Feinberg suggests that at least some human beings should be excluded from this group and wonders whether or not this would be too heavy a price to pay.

Another view is that a being must have interests if it is to have any rights; in other words, it must care, or be concerned about what happens to it. Not surprisingly, this has again led to a widespread debate. Some thinkers like Feinberg[8] have argued that animals have a 'conative life' with wishes, hopes, urges and impulses, thus raising the possibility of 'the good' of an individual animal. Other authors such as McCloskey[9] have consistently opposed such a view since they are of the opinion that animals have no demonstrable interests.

Yet another property of a being with rights has been believed to be sentience, that is, the capacity to feel pain and pleasure, a point that has again been discussed above. Ritchie[11] makes a very strong argument against attributing rights to animals on this ground. He notes that animals can never be accorded equal rights with human beings since this would imply, rather absurdly, that no animal could be put to death without a fair trial, and that it would be our duty to protect weak animals from the attacks of the strong. If, on the other hand, they are considered to have rights inferior to those of man, this would merely be an excuse for human interests to override the interests of animals in determining how animals themselves should be treated.

It should, of course, be made clear that even if we assume that animals are beings that can potentially have rights, we still have no answer to the question of whether they actually do have rights. Another dimension to this debate is added by thinkers who believe that animals do have rights but seem to disagree on what exactly are the rights that animals possess. Regan[12], for example, believes that they have a natural right to life if it is supposed that all human beings do, while Feinberg[8] feels that they have a right not to be treated cruelly. Salt[10] believes that animals have a right to exercise their cognitive and emotional capacities, and Rachels[13] argues that at least some animals have the right to property and to liberty. This point is a particularly interesting one since Rachels clearly refers to moral freedom exactly in the same way that has been proposed for human beings as well, and recent experiments with chimpanzees have shown that they have limited abilities to express empathy and compassion, human traits traditionally regarded to be morally admirable[14].

Where do we stand?

The problem of animal rights has obviously raised more heat than shed light on what our obligations should be towards all nonhuman animals, particularly in the way we treat them for our benefit. Some of the contentious issues relate to what is a right, especially for an animal, and what kinds of natural rights can animals lay claim to while being entitled to other special rights. Other questions, such as whether animals do have interests, or how self-aware or conscious they are, are more factual and may even be answerable one day given our new interest in animal cognition and our enquiry into the workings of the animal mind.

The whole issue, however, in my mind, still remains as ambiguous and controversial as it was in the day of Aristotle. And the foremost reason for this seems to be the inevitable subjectivity with which each of us perceives animals and our relations with them. We still

On Shri Morarji Desai and the rhesus macaque

'The moral status of a country is known by the way it treats its animals.'

– Mahatma Gandhi

The rhesus macaque (*Macaca mulatta*) of north India has, since ancient times, played an important role in the ecology, culture and traditions of India. The close interaction of these macaques with the people of our country forms perhaps the most intense relationship between human and nonhuman primates anywhere in the world.

Easily maintained under laboratory conditions, however, the rhesus monkey has been extensively exploited in biomedical research because its disease spectrum is apparently very similar to that of human beings. In the early sixties, India was exporting about 200,000 to 300,000 monkeys *every year* to Europe and USA, and this had obviously considerable impact on the wild populations of this species.

In 1977, the International Primate Protection League (IPPL) gathered documents about the apparent misuse of imported Indian rhesus monkeys in military experimentation in the United States. India had lifted an initial ban on rhesus exports in the late fifties. This ban had resulted from the heavy mortality occurring amongst the transported monkeys. There were, however, conditions placed on the use of monkeys when exports resumed. Each individual shipment had to be accompanied by a 'certificate of need' which guaranteed '. . . (1) that the monkeys now being purchased will be used only for medical research or the production of antipoliomyelitis vaccine . . . ' and (2) that . . . 'they will receive humane treatment under our care'.

The IPPL felt that the radiation experiments being performed on hundreds of rhesus monkeys by the US military at various installations flagrantly violated the US–India agreement: in some experiments, monkeys were trained by electric shock to run in large circular treadwheels, they were then exposed to massive lethal doses of radiation and placed back in the treadwheel, and their agonized efforts to run were observed by scientists. In one experiment involving 131 monkeys exposed to between 2500 and 80,000 rads, the following responses to radiation were seen:

'Extreme forward slumping in crouching posture; whole body convulsions; spasticity; rolling eyeballs; stumbling and falling; blind, apathetic facial expression; muscular tremors; vomiting and retching; frequent shifts in body position and posture; wild purposeless twisting, twirling and throwing of the body; uncoordinated leaps on hearing a sound; crawling and frantic pushing and scraping of sides, back and belly on floor of cage; and passive draping of the head over the back while in an awkward, crouched, clinging posture.'

The IPPL initially tried to get the US authorities to stop these experiments. On failing to do so, they contacted the then Indian Prime Minister, the late Shri Morarji Desai, their Indian advisers and the Indian Press in order to mobilize popular and government opinion.

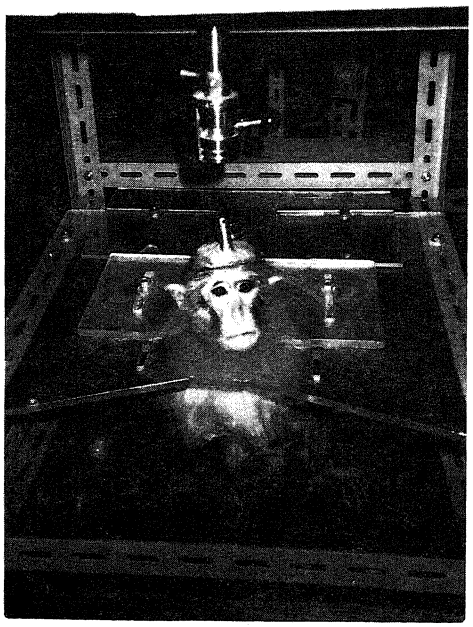
On 3 December 1977, India's export ban on monkeys was officially announced. It is to the credit of Shri Desai that in the face of increasingly mounting pressure from powerful lobbies, he stood firm. This policy was later continued by his successors including the late Smt Indira Gandhi, again inspite of considerable political pressure.

In an attempt at historical revisionism, claims were made by US scientists (in a magazine article) that the Indian ban had resulted from conservation concerns and the dwindling number of the rhesus. This claim was subsequently retracted when Shri Desai, who had by then retired, countered this in a handwritten letter to the IPPL:

'You are quite correct in saying that I banned the export of monkeys on a humanitarian basis and not because the number was lessening. I believe in preventing cruelty to all living beings in any form. This is the ancient Indian culture . . .'

Later, a survey by the Zoological Survey of India determined that there were only 200,000 rhesus monkeys left in India. The trade had taken a heavy toll. The teeming millions of former days had indeed disappeared.

A life-long vegetarian and animal lover, Shri Morarji Desai rejected medical treatment and vaccinations throughout his life. The monkeys left may not know it, but they owe their lives and freedom to him. They are, in their way, his living monument.



A rhesus macaque confined to a restraint chair with electrodes implanted in its brain during the course of an electrophysiological experiment. Photograph courtesy: The International Primate Protection League.

have not been able to come to terms with our biases and prejudices in the matter of human race relations, religion and sexuality. What hope remains for our attitude towards animals?

Schweitzer[6] fully grasped the enormity of this problem and realized that we are indeed far from a final solution. He thus concludes his brilliant essay entitled 'The ethic of reverence for life' with the following words, and I can think of no better way than to end this essay with them:

'Wherever any animal is forced into the service of man, the sufferings that it has to bear on that account are the concern of every one of us. No one ought to permit, insofar as he can prevent it, pain or suffering for which he will not take the responsibility . . . The ethic of reverence for life . . . inspires us to join in a search for opportunities to afford help of some kind or other to the animals, to make up for the great amount of misery which they endure at out hands, and thus to escape for the moment from the inconceivable horrors of existence.'[6]

1. Reported in *Science Digest*, November 1973, p. 32.
2. Regan, T. and Singer, P. (eds), *Animal Rights and Human Obligations*, Prentice-Hall, New Jersey, 1976.
3. Aquinas, St. Thomas, *Summa Contra Gentiles*, (translated by the English Dominican Fathers), Benziger Brothers, 1928. Reprinted in *Animal Rights and Human Obligations* (eds Regan, T. and Singer, P.), Prentice-Hall, New Jersey, 1976, pp. 118–121.
4. Plutarch, reprinted in *Animal Rights and Human Obligations* (eds. Regan, T. and Singer, P.), Prentice-Hall, New Jersey, 1976, pp. 111–117.
5. Bentham, J., *The Principles of Morals and Legislation*, 1789. Reprinted in *Animal Rights and Human Obligations* (eds Regan, T. and Singer, P.), Prentice-Hall, New Jersey, 1976, pp. 129–130.
6. Schweitzer, A., in *Civilization and Ethics* (*The Philosophy of Civilization*, Part II), Macmillan, New York, 1950, pp. 130–138.
7. Singer, P., in *Animal Rights and Human Obligations* (eds Regan, T. and Singer, P.), Prentice-Hall, New Jersey, 1976, pp. 148–162.
8. Feinberg, J., in *Philosophy and Environmental Crisis* (ed. Blackstone, W.), University of Georgia Press, Georgia, 1974, pp. 190–196.
9. McCloskey, H. J., *Phil. Quart.*, 1965, **15**, 115–127.
10. Salt, H. S., *Animals' Rights*, The Humanitarian League, London, 1912.
11. Ritchie, D. G., *Natural Rights*, Allen and Unwin, London, 1894.
12. Regan, T., reprinted in *Animal Rights and Human Obligations* (eds Regan, T. and Singer, P.), Prentice-Hall, New Jersey, 1976, pp. 197–204.
13. Rachels, J., in *Animal Rights and Human Obligations* (eds Regan, T. and Singer, P.), Prentice-Hall, New Jersey, 1976, pp. 205–223.
14. Cheney, D. L. and Seyfarth, R. M., *How Monkeys See the World*, University of Chicago Press, Chicago, 1990.

In late 1994, Jane Goodall, the eminent primatologist and now a noted advocate for the noninvasive use of animals in research, gave a call to academicians the world over to participate in a dialogue. This was essentially aimed not only at raising the awareness of researchers, but more importantly, at opening up a serious discussion on the ethics of using laboratory animals in often painful and sometimes relatively unwarranted basic and medical research. We reprint below some excerpts from this call.

– Editors

On the use of animals in research and education

Jane Goodall

Green World Center, PO Box 45, Highgate Springs, Vermont 05460, USA

The use of nonhuman animals for the purposes of humans has long been taken for granted in our culture, and has been institutionalized by entire industries. In recent years, however, a new awareness of animals has been developing, and new attitudes and practices have come into being.

Over the last two decades, the ethical and broad scientific implications of the use of animals in laboratory experiments have come to be examined more and more critically, and new research methods have been developed. There may now be some consensus among scientists, as well as among the public, that the use of animals raises ethical questions that must be dealt with.

However, discussion of the use of animals in general and as experimental subjects in particular has been polarized and contentious. Research scientists and animal rights advocates have regarded each other with distrust, and constructive dialogue has been scarce. Stereotypes of the researcher as unfeeling, and of the animal advocate as fanatical, have been persistent.

It is time for the ethical, scientific and practical issues raised by the use of animals in research and education to be aired anew, with a fresh measure of good will.

By and large, institutions of higher learning have not paid enough attention to the status and treatment of animals in society and in the institutions themselves. We,



Animals are frequently subjected to avoidable torture. This animal is fistulated – a process of cutting a circular hole of about 25 cm and covering with a removable plug so as to draw the contents from inside whenever needed.

therefore, urge that educational programmes and public forums, such as conferences, symposia, debates, film showings and seminars, be organized to consider the full range of questions raised by animal experimentation and by the use of animals in academic settings.

For example:

♦ What are the ethics of using animals as tools for human purposes, and of invasive or otherwise harmful experimentation on animals? Do animals have a right to be treated as ends in themselves and not as means only?

♦ What are the practical limits on the pain and suffering to which nonhuman animals in laboratories can be subjected? What kind of protection is given by existing laws and institutions? How common are practices that involve unalleviated pain? Are such experiments ever justified? Is there a conflict between limits on pain and suffering and the advancement of science or academic freedom?

♦ What animals are subjected to what experiments at your school? What are the experiments most harmful to animals? Do any involve causing pain without administering pain killers?

♦ To what degree have animal experiments been useful and valuable? What role have they played in the history of medicine? With respect to the future: would human health suffer if medical research were guided by a new vision in which animal experiments had no place? Would it benefit?

♦ There is an array of research methods, including clinical study, epidemiology, *in vitro* tissue culture research and computer modelling, that do not use animals. What is the adequacy and scope of these methods?

♦ Animal experiments generally begin not with sick animals in need of healing, but with healthy animals, who may then be innoculated with approximations of human diseases, or exposed to toxic substances, or subjected to a variety of surgical or other procedures. Psychological experiments have included work on such themes as 'maternal deprivation' of primates and 'learned helplessness', and military experimenters subject animals to wounds and burns. What would be the effect on medical science if experiments were limited to attempts to heal already sick or injured animals?

♦ Is the elimination of animal experiments a worthwhile goal? What about reduction in numbers of animals used, replacement of animals with nonanimal methods, and 'refinement' of experimental procedures? If such goals are worthwhile, how should they best be achieved? What is being done towards these ends today, and what is not being done that could be? How much room for improvement is there?

♦ Under what conditions are the animals at your school housed and cared for? Are their cages kept clean? Are their physical needs cared for adequately? Do they have veterinary care? Are their psychological and social needs met? How does the disruption of normal family life affect them? Do they receive loving attention? Can they be visited by members of the community? If not, why not?

♦ Millions of animals each year are killed for dissection to teach students life sciences. Is dissection necessary as an educational tool? What options are in use at your school for students who conscientiously object to dissection?

♦ Who are the suppliers of animals for dissection at your school? How many animals are killed each year for dissection? How much is paid for them? What are the conditions under which these animals are bred or captured and under which they are maintained and killed? Are visits to the suppliers' premises possible?

***

For help organizing a forum, The Green World Center will provide consultation, advice, and resources concerning

(a) organizing forums, film showings, educational activities and events;

(b) educating yourself and others about the issues related to animal experimentation;

(c) learning about animal experimentation on your campus.

Resources include speakers, literature, and an excellent Canadian Broadcasting Company video on animal research with David Suzuki.

If you think you might like to help in any capacity, or if you just want to learn more about animal experimentation, please write to or phone:

Walter Miale
Green World Center
PO Box 45, Highgate Springs, Vermont 05460, USA
C.P. 29, Philipsburg (Qubec) J0J 1N0, Canada
Phone and fax: 514-248-7575; if busy: 514-2483868
e-mail: F1C4@MUSICB.McGill.CA

Human–animal interactions

N. K. Waran

Institute of Ecology and Resource Management, School of Agriculture, University of Edinburgh, West Mains Road, Edinburgh EH9 3J/G, Scotland, UK

There has been limited research conducted on the importance of the human–animal relationship. This paper describes the ways in which human behaviour or attitude might influence the welfare of tamed animals.

That man is important in an animal's environment is not disputed. Direct interactions between humans and animals have always been a part of the animal production system. The way in which the animals respond to human contact is governed by the way in which they perceive humans, and this is influenced both by their early developmental experience and by their natural responses or genetics. In general, it is believed that domestic animals are less fearful of humans, their flight distances are shorter than their wild counterparts, and they are more able to cope with human intervention. But animal responses are at least in part governed by fear. The fear response of animals is thought to be one cause of a chronic stress situation. Simply by being in, and imposing themselves on, the environment of an animal, humans induce fear and adverse reactions in the animals[1].

Animal responses to humans have been measured. For example, in battery-housed hens there was an increase in the resting heart rate from 140–240 beats/min up to 280–350 beats/min when a human entered the building; and rose to 480 beats/min when the human approached the cage and encroached upon the birds' natural flight zone (or personal space).

Effect of handling animals

Recent studies have shown that different handling procedures influence the performance and behaviour of animals. Growth rates in pigs were shown to be adversely influenced by what were termed 'aversive' and 'negative' handling. Pigs were conditioned to elicit a particular response from a human every time they interacted. The aversive treatment involved prodding with an electric prod every time a pig approached a human, and negative treatments involved a human approaching the pig in an upright posture and reaching towards it. Growth rates were approximately 10% lower in these treatments than in minimal contact or positive contact treatments. The reduction in growth rate was attributed to the chronic psychological stress experienced by the animal as a consequence of the handling treatments. In order to test this, the morphology of each treatment groups' adrenal glands was examined; it was found that the area of the adrenal cortex in pigs that experienced aversive treatments was almost 30% greater than those subjected to minimum, or positive treatments[2].

In other studies[3], reproductive performance has been shown to be influenced by handling treatments. When gilts were conditioned to expect to be scratched and stroked by humans, or pleasantly treated, they were easily mated at second oestrus, with a pregnancy rate of 87.5% as compared to only 33% in gilts that had been conditioned to expect a shock from an electric prodder. Boars in the unpleasant treatment groups were on an average 31 days older before they demonstrated a fully coordinated mating response, and their testicles were 10 cm^2 smaller. Measures of the concentration of plasma cortiosteroids in the absence of humans showed higher concentrations in all pigs following the unpleasant treatment.

Thus, it appears that different forms of handling can influence growth and productivity, and that it is probable that this is due to the animals' fear response to humans. There are several more examples of the effect humans have on farm animals: milk yield has been shown to vary between stockpeople; handling can negatively affect ovulation rate in ewes; regular handling during early development can enhance growth in broilers (neat-producing hens), and so on. It is obvious that

we need to understand more about these effects and, in particular, we need to know how to have a positive effect on the animals under our care.

Is there an optimum level of handling?

It has been suggested that one way to influence an animal's perception of humans is to handle in a positive way early in development. This first indication of the positive effects of handling in early development was shown in Denenberg's work on the 'gentling' of rats in 1964 (ref. 4). Young rats were stroked for a short time every day, what Denenberg termed 'introducing extrinsic stimulation'. He found that rats performed better in learning experiments and were less fearful in tests later. Although in most cases this early handling has been shown to reduce the fear of humans in rabbits, chickens, pigs and cows[5], it is often a very small effect and tends to be specific to certain handling procedures. In some cases, the effect of handling early on has had no effect, or even a negative effect, making animals even more fearful. For example, when calves were artificially groomed in the first 24 h of life, they stood up faster than ungroomed controls but they absorbed immunoglobulins at a lower rate[6]. There is very little understanding of the mechanisms underlying these apparent positive effects. Handling pigs at an early age can reduce excitability, making them easier to drive[6], but too much handling may result in an animal with such a reduced flight zone that it no longer moves away from a human, thus making it impossible to handle. A good example of this is orphan lambs reared by humans – apart from the health problems they may experience, these animals are often difficult to manœuvre as part of a flock.

It would thus appear that there is an optimum level of handling during early development that leads to a reduction in fear responses toward humans, but leaves the human the dominant figure in the animal's environment.

Are there sensitive periods of response to handling?

It has been suggested that humans can take advantage of the sensitive periods during an animal's development to condition it to certain cues. This is most obvious in birds that are reported to have sensitive or critical periods after hatching, when the bird is particularly sensitive to extrinsic stimulation. The term 'imprinting' describes the attachment that birds develop to their parents or peers during this time. In a study of game birds, where it is important that birds are fearful of humans, it was found that exposure to humans during the first 2 days after hatching resulted in birds being less fearful of humans, allowing them to approach much closer before showing the escape behaviour than birds exposed to humans only after the first 2 days. It has been suggested that other animals are also particularly sensitive at certain times in development. Hemsworth and his colleagues suggest that a sensitive period existed in pigs prior to 8 weeks of age, when they could be socialized to humans[5]. Sensitive periods may occur also later on in life, such as during parturition, when bonding between dam and offspring takes place. Cows handled during their first calving have been shown to be easier to handle and display less flinch, step and kick responses during their first 2 weeks of milking, indicating that they were less fearful of the handling procedure. It has also been suggested that by encouraging the cows to develop a general attachment to humans during the sensitive period at parturition, the cows become less fearful later[6].

Training the animals

Animals can also be trained or conditioned to accept certain restraining devices if they are positively rewarded in some way. Teaching animals to behave in an appropriate way depends on the animals' ability to make certain associations, though animals are thought to be generally limited in their ability to learn. Constraints on learning have been seen in laboratory experiments with pigs which, in response to a food reward, tend to perform their species-typical food-searching activity – rooting[8]. In addition, training depends on an animal being able to remember the association. Training is often time-consuming and sometimes impractical where animals remain in a system for only a short time; but it has been used with sheep and cows with promising results. Sheep trained to run a race and to enter a handling machine for barley food rewards, have been shown[9] to run the race quickly from year to year. Dairy cows have been conditioned to move from their field to the milking parlour in response to the sound of a bell, and an estimated 1 h of a stockman's day can be saved in this way[10]. Natural leader sheep or Judas sheep can be trained to lead a flock through a slaughterhouse[11].

It appears, therefore, that humans can have considerable positive effects upon farm animals but can also have negative effects. Human behaviour is often inconsistent and dependent upon mood, and humans are often perceived as a fear-provoking stimulus. It has been suggested that automation of animal units might be more predictable and, therefore, more easy to 'adapt' to. Mechanized broiler and turkey catchers are currently being developed and have been shown to provoke lower stress responses than normal procedures by humans. Food for thought!

1. Metz, J. H. M., in *The Role of the Stockman in Livestock Production and Management* (ed. Seabrook, M. F.), CEC (EUR 10982), Luxembourg, 1987.

2. Gonyon, H. W., Hemsworth, P. H. and Barnett, J. L., *Appl. Anim. Behav. Sci.*, 1986, **16**, 269–278.
3. Hemsworth, P. H., Barnett, J. L. and Hansen, C., *Appl. Anim. Behav. Sci.*, 1986, **16**, 259–267.
4. Denenberg, V. H., *Psychol. Rev.*, 1964, **71**, 335–351.
5. Hemsworth, P. H., Barnett, J. L. and Coleman, G. J., *Anim. Welfare*, 1993, **2**, 33–51.
6. Grandin, T., *Lab Anim.*, 1986, **15**.
7. Csermely, S., Mainardi, S. and Spano, S., *Appl. Anim. Ethol.*, 1983, **11**, 177–182.
8. Breland, K. and Breland, M., *Am. Psychol.*, 1961, **16**, 661–664.
9. Hutson, *Appl. Anim. Behav Sci.*, 1985, **15**, 263.
10. Kiley-Worthington, M. and Savage, P., *Appl. Anim. Ethol.*, 1978, **4**, 119–124.
11. Brenner, K., *NZ J. Agric.*, 1980, September, 25.

Conflicts or coexistence?

Madhusudan Katti
Department of Biology, University of California, San Diego, USA

The sprawling American city of Los Angeles has obliterated vast areas of natural habitat in southern California, and its suburbs continue to encroach upon the remaining fragments of forest and coastal sage scrub, for condominiums and recreation. The Angelenos forest has, amazingly enough, survived to this day, and is one of the remnant fragments of wilderness near the city. Even this is surrounded by suburbia – and is a source of frequent conflicts between humans and wildlife.

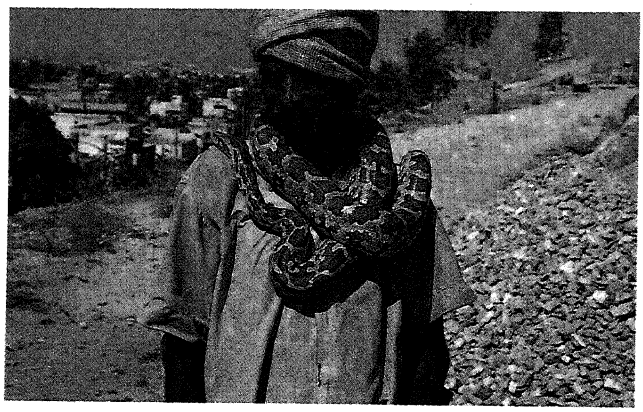
A young mother jogging alone one evening on a mountain trail in the suburbs got killed by a mountain lion: the first known mountain lion attack on humans this century! The local police and fish and wildlife officials immediately swung into action and, in a couple of days, tracked down and killed a female mountain lion. Shortly after they confirmed – using genetic tests of hair samples – that it was indeed the same killer animal, a solitary mountain lion cub was found near where the jogger was attacked. The cub was then sent to a zoo. It seems likely that the jogger strayed too close to the cub – *Had she seen the animals? Did she deliberately approach them too close (like so many Americans brought up on the ideas of cuddly wild animals often do)? Or did she just blunder into them*? – and provoked the mother into attacking her.

About a month after the mountain lion attack, a big black bear strayed out of the Angelenos forest and wandered into a suburban mall. Panic ensued and the local police and wildlife warden were alerted. The game warden soon shot the bear dead after unsuccessful attempts to tranquilize it for translocation. A few weeks later, a panic-stricken 911 (the emergency telephone line) call from another LA suburb reported another black bear which the caller could see outside her/his window near a school. When the cops and game warden rushed in, they found the bear perched up in a tree *looking at some children playing in the schoolyard* – so it had to be shot dead too!

Reading about these incidents in the long, hot Californian summer of 1994, when a vast forested area outside LA was burnt down (by arsonists?) and huge tracts continued to burn all over the northwestern US, I could not help but wonder if we in the third world had anything to learn at all from this, the world's most advanced, industrialized, first-world nation, in terms of managing our wildlife and natural habitats.

I myself grew up in Bombay, a city often compared to Los Angeles or New York, but much worse in terms of infrastructure or civic amenities, and on an average, offering a lower quality of life to its inhabitants. Unlike LA, Bombay also faces a shortage of land to accommodate its burgeoning human population. Yet, bordering the megalopolis' northern suburb of Borivli lies a 100 km^2 tract of tropical forest protected in the form of a National Park. The forest is in remarkably good shape, considering it is completely surrounded by over 15 million human beings. The park also has a resident population of leopards estimated at around 20–40: certainly too high a density of large carnivores to be supported by the habitat. The Indian leopard is a larger animal than the North American mountain lion, with a long and continuing history of man-eating problems that keep cropping up all over the country each year. Many of the Borivli leopards, lacking space and food inside the park, venture outside the forest at dusk to pick up dogs and other small animals from the shanty towns bordering the park. Sometimes the prey include human children, and even adult humans have been attacked and killed. Yet, the leopards are tolerated and their population allowed to build up over the past many years. It was only recently – responding to a public outcry after a child's killing – that the forest department trapped two individuals, which, unfortunately enough, bashed themselves against the cages and died.

The Borivli leopards are by no means an exception to man–animal relations in India. There is another long-standing, recurring conflict between a large mammal and human beings in eastern India for which the solution still eludes the Forest Department. This is the case of the famous Dalma or Bihar – Bengal herd of elephants:



Animals are often entwined into the lives of rural Indians. Although frequently employed in earning a livelihood, they are also treated as members of the extended family.

a remnant population of about 40 animals that has lost almost its entire home range of tropical deciduous forests to the axe and the plough. This herd – of a species renowned for its long memory – flow wanders across the hills and plains as if in search of its lost homeland. Inevitably, they come into conflict with peasant humanity, which happens to produce the only foods now available to the elephants. So the elephants undertake their annual migrations after the monsoons carving a trail of destruction among the farms and villages: large areas of crops are laid waste – from the human point of view – and frequently, humans who dare to defy the giants get killed. Last year, its incredible march led this herd almost into Calcutta city: the Forest Department and villagers only barely managed to turn the elephants back from some 20 km outside the city, averting what could have been a disaster of spectacularly tragic proportions. Once again, in a now annual feat, the foresters and villagers managed to beat them back to their patch of remaining forest sanctuary – only to gain a few months of relief before the elephants set forth once more. As with the Bombay leopards, the only action the government has taken so far is to pay some compensation to the victims and talk of translocating the herd. Meanwhile, the people are already taking things into their own hands: at least three elephants were killed last year – victims of poisoning or electrified fences.

Virtually all the remaining wild elephants in India face this problem of habitat loss and fragmentation to some extent, and conflicts with humans are common. One recent case was that of a big tusker which took to raiding crops outside the Kalakad–Mundanthurai Tiger Reserve in Tamil Nadu and killed two villagers before it was itself tragically brought down by a fatal dose of tranquilizers pumped into it in an attempt to capture and translocate it. But there was little else the beleaguered Forest Department could have done – if they had not made the capture attempt, people may have taken things into their own hands with graver consequences. And so the casualties mount on both sides of this conflict.

Going back to the question of whether the West can offer us any new ideas on dealing with the problem of wildlife in fragments of natural habitat embedded in a humanized landscape matrix: the above stories seem to highlight a certain contrast between the two societies' approaches. On the one hand, we have a country of 950 million people, most of whom lead a hand-to-mouth existence, with hardly enough land to support them – but willing to share that land with its wild denizens, even at a substantial human cost. On the other hand is a country with barely a fourth of India's population occupying a landmass three times as big as India, and a much higher standard of living for its people – yet even the smallest trespass by a wild animals brings instant and fatal retribution! The way many western countries have systematically targeted wild animal species perceived to be inimical to human interests (largely economical), often with little real evidence – like the wolf and the coyote in North America and the dingo in Australia – suggests a basic lack of respect for Nature and its wild denizens. This contrast is indeed a historical one: one can certainly make a strong case that Indian society has always been more protective of its wildlife than any western nation, the vast differences in population, resources and standards of living notwithstanding.

Some of my American teachers, while lecturing students on the loss of biodiversity and human pressures on wildlife, like to show a clipping from a film on human population growth. The segment shows a map of the world with a number of red dots scattered across it: the animated map depicts the distribution of human population across the globe, each red dot representing a population centre of over 1 million people. As the animation starts from a few thousand years ago and reaches the present time, one sees the map filling up with red dots: initially quite slowly, but at an accelerating pace in the current century. The film makes quite a dramatic illustration of human population explosion and, accompanied with data on global extinction rates of different species, reinforces the popular correlation between population growth and loss of biodiversity. The subtext is the fact that the equatorial belt, or the South, with the maximum number of red dots is also the region with greatest threats to biodiversity.

But what struck me most was another element which actually subverts the above popular analysis: the fact that the Indian subcontinent – as indeed most of Asia – has always supported a disproportionately larger share of red dots. Yet the extinction rates in this part of the world did not shoot up until the arrival of Europe in India! Another interesting region worthy of such a fine analysis is South America: that continent had a fairly high density of red dots before Columbus discovered it, but many of these blinked out following the arrival of

Europeans. And again, there is little evidence suggesting extinctions caused by native Americans: the overall positive correlation between human population growth and extinction rates appears here too only in the post-Columbus era! This pattern has already been pointed out in some critiques of the biological expansion of Europe – but it remains in the realm of the subversive, on the fringes of mainstream ecological discourse, and has not made it into many university syllabi.

This suggests that the differences are indeed cultural, with the less modern pre-western societies being more tolerant of wildlife than the modern western ones. And so, one may be reassured about our own historically superior record – at least in this respect – and search for the older ways that allowed us to maintain this coexistence. But I suspect the Indian scenario is (typically, one may say) much more complex, much more ambiguous, given the overt and subtle interplay of cultural, social and economic differences inherent in our multilayered society. If the above examples highlight coexistence in Indian society at a broader level, the following stories suggest streaks of arrogant intolerance at some specific sublevels.

The story of the Sunderban – deprived humans pitted against man-eating tigers in the world's largest mangrove swamp and tiger reserve – is all too sordidly familiar. Perhaps the final solution – one way or the other – to this intractable conflict will come from global warming and its attendant rise in sea levels! And that gives rise to a vision of the homeless humans of Calcutta stranded between man-eating tigers on one side and elephants on the other, both rendered homeless due to different human activities – one shudders at such an image! But tigers do occasionally turn to human flesh in other areas as well, even if not on the same scale as in Sunderban: while the much-lauded Project Tiger has so far succeeded in protecting many tiger populations, it has failed to provide more habitat for the increasing tiger population, and individuals are often forced to forage outside the reserves. Man-eating problems are not infrequent in some of the sub-Himalayan tiger reserves, notably Dudhwa, and the Forest Department has often been forced to kill some of these problem animals. One particular instance from the late 1980s, however, is stuck in my mind. One tiger had, over the course of a few months, killed as many as 16 local villagers, but no action was taken: perhaps the villagers were trespassing to collect firewood or engaged in another such illegal act, and so were either not vocal enough in their protests, or otherwise invisible to the authorities. But then the tiger made the error of picking its 17th victim from among the staff of the Forest Department – and was killed soon after. As one of my teachers at the Wildlife Institute of India commented wryly, it may seem that in some parts of India the life of one forest official equals those of 16 poor villagers!

Going further back, even if it was the arrival of the Europeans in India that triggered off major ecological damage, these external forces were, nonetheless, aided and abetted by many within our society. While the white hunter may have brought the guns, it was largely the local Maharajas, Princes and Nawabs who drove the cheetah to extinction and pushed the tiger and many other species to the brink. On the other side of the human/wildlife conflict/coexistence coin are the Bishnois of Rajasthan: now an enduring symbol of a traditional conservation ethic with great respect for wildlife, hundreds of Bishnois were massacred 300 years ago for protecting the trees around their villages – not by any European imperialist agency but by their own local king.

Where does that leave us then? On the one hand, we show higher tolerance for wildlife than the West, even bearing some human costs. But on the other, the West treats its citizens more as equals when it comes to wildlife conflicts – each victim is avenged, so to speak – while we seem to place our own in a hierarchy of values. This might lead one to say that in India wildlife has more value than (some) human life – a position many of the poorer victims already articulate. Perhaps, it is not too surprising to learn that in a society where some people have always been more equal than others, some animals might also be held more equal than some humans! The West, though it has its human hierarchies, is nevertheless more explicitly homocentric when it comes to dealing with nonhuman nature.

What lessons may one draw from such comparisons? Neither approach satisfies anyone who is interested in ensuring justice for both the human and nonhuman life forms sharing this earth. The Indian approach of tolerance and sharing allows nonhumans to coexist with humans to a greater extent – but many humans suffer. Even these older values and ethics respecting nature are now retreating under the combined assault of western consumerism and increasing human needs within the country. The West is still in the process of evolving a new set of ecological ethics based on a modern biological understanding of the origin and role of life on earth. We may be in a position to offer a model (religion-based) framework for such ethics, which can be expanded and elaborated with the help of modern ecological understanding. It is perhaps time we start actively offering the West some of our knowledge and solutions to its own problems, rather than passively waiting for them to come and help solve ours. For neither culture knows the best answers and both are groping for the light – albeit one more confidently and assertively than the other.

Environmental design

N. K. Waran

Institute of Ecology and Resource Management, School of Agriculture, University of Edinburgh, West Mains Road, Edinburgh EH9 3JG, Scotland, UK

Investigations of the behaviour of farmed animals kept in intensive conditions in Europe and America have indicated that their welfare is generally of concern. The inherent features of the modern systems for housing animals are confinement and restriction. Today's husbandry systems appear to be problematic for animals in that they prevent them from behaving in the way intended by nature. Consider, for example, the plight of the pig, a highly exploratory and socially sophisticated animal. When the sow gives birth, it is generally in a crate that prevents her from squashing her piglets. It also prevents her from building a nest, something she is highly motivated to do prior to parturition, and she is prevented from moving about, turning around and interacting normally with her piglets. Piglets are weaned at an unnaturally early age (3–5 weeks) and are moved into weaner accommodation that is typically cramped (0.2 m^2/pig) and barren (no bedding, no rooting substrate). They remain in grower accommodation until they reach slaughter weight, when they are often mixed with unfamiliar pigs and transported in cramped conditions to the abattoir.

Many of the animals in intensive systems exhibit abnormal behaviours which are sometimes stereotypic[1]. Stereotypies are defined as behaviours that are performed with monotonous regularity and are fixed in all details. Some researchers believe that they are mechanisms enabling the animal to cope with an environment that does not satisfy its expectations[2]. Others believe that they are associated with levels of feeding; pigs that are confined and fed low-maintenance diets are more likely to exhibit such behaviours than those fed higher levels. What does seem clear is that animals do not show these behaviours if they are kept in husbandry systems that are similar to the natural environment of their species.

In general, animal housing in the western world has evolved to suit the requirements of humans rather than of animals. The importance of the ease of cleaning and managing housing, and of handling, moving and observing animals, has quite correctly been catered for in most environments; but the actual behavioural needs of animals appear to have been compromised in many cases. Certain natural behaviours are not seen in the modern housing system, and more often than not animals perform abnormal or stereotypic behaviours. Behaviours such as feather-pecking in battery hens, belly-nosing in weaner pigs, and bar-biting in stalled sows are thought to be indicative of welfare problems in these animals, and are probably caused for the most part by the conditions under which they are kept.

Recently there has been a growing concern about the welfare of animals in intensive conditions, and with this, a growing interest in the design of so-called *ethologically sound* environments: environments that facilitate the species' typical behaviour and also enable that animal to be farmed effectively. There have been two broad approaches used in the past in designing environments that suit the behavioural needs of animals, or ethologically sound environments. The most well-known approach, and innovative at the time, was that used by the late David Wood-Gush and Alex Stolba (of Edinburgh, Scotland) and others in the designing of housing for pigs. This 'start-from-scratch' approach enabled the researchers to identify key features in the pigs' normal environment that appeared to elicit or release certain behaviours[3]. These key features were then incorporated into a new housing system for pigs called the Edinburgh Family Pig Pens. The second approach is one of modifying an existing environment through the improvement of the quality of limited space, either by artificial enrichment devices (such as food-balls and hanging tyres/chains for pigs) or by incorporating certain features thought to be important to the animals (such as dust-baths and laying-boxes for hens)[4–6].

Start-from-scratch designs

In general, it is often suggested that domestic animals have been heavily selected for certain features of domestication (for example, docility, trainability and a reduced fear of novel situations) and have 'adapted' in ways that suit the modern farming systems; in the process, they have lost the ability and the need to exhibit the behaviours their ancestors would have shown in their natural environments. In order to test this theory, in 1978, David Wood-Gush, Alex Stolba and colleagues released ordinary domestic large white sows and boars in a semi-natural environment in the Pentland Hills outside Edinburgh[3]. The Edinburgh Pig Park, as it was called, consisted of a 1.1 hectare site on the side of a hill which had forested and open areas, different types of grasses and shrubs, and various degrees of slopes. The group composition of 2–5 sows and 1–2 boars plus



Animal housing in the western world has evolved to suit the requirements of humans rather than of animals.... Certain natural behaviours are not seen in the modern housing system.

1–10 juveniles was adhered to, and the pigs were allowed to live and breed as normally as possible. Over a number of years, various studies on the behaviour of sows and piglets were conducted. These studies suggested, contrary to the opinion at the time, that pigs exhibited most, if not all, of the behaviours that had been recorded in groups of wild pigs. It was possible to identify certain key features in the pigs' environment that were important in the release of certain behaviours, such as rooting (i.e. digging with the snout) and defecation sites, and nesting sites. Based on this Stolba and Wood-Gush[3] attempted to recreate the 'natural' environment in a more restricted indoor area that could easily be managed. Although their aim was to develop an environment that allowed pigs to exhibit more natural behaviours, and to live in a more natural grouping, they also wanted to make this alternative style of housing more economically viable.

In 1980 the first Edinburgh Family Pig Pen was built at Easter Howgate Farm, south of Edinburgh. The Pig Pen housed four sows and their offspring. The sows remained there until all their breeding life, and were joined by a familiar boar 2–3 weeks after farrowing. Subsequent study of this system resulted in alterations in the original design, and a further two layouts were tested. Despite these alterations, there were still problems with the design. The monopitch roof of the building coupled with the cold Scottish weather was blamed for the high incidences of piglet mortality, and it was difficult to maintain an optimum group size of four sows since the addition of new individuals to the established group caused problems. However, the system did appear at the time to be economically viable – the number of piglets weaned was comparable with the better intensive systems, food conversion and growth were also comparable, and production records were good. The system did not take off commercially but the approach highlighted the importance of certain features for pigs, and showed that pigs could be kept in groups. The system is still being refined to this day: A research group in Aberdeen (Scotland) is working on a pen that operates on these lines, and a group in Switzerland (the home of the late Alex Stolba) is also developing the system.

There are, however, additional problems in this approach. Although the pigs were seen to exhibit certain behaviours in a more natural setting, no attempt was made to determine the importance of these behaviours for the animal. The need to show a behaviour, or the strength of the preference for an environmental feature, was not measured at the time, and the fact that the animals exhibited certain behaviours when given the opportunity to do so was taken by the researchers to indicate that these were important for the animals' well-being. This need not be the case since some behaviours may be considered as 'luxury behaviour': the animal may show it if it can but is not highly motivated to perform it if it cannot.

Recently, methods of assessing an animal's need to exhibit a behaviour have been developed[7]. If these methods are used in conjunction with the 'start-from-scratch' approach, a clearer idea of the importance of each of the key features may be gained.

Altering the existing designs

The second approach, that of modification or enrichment, involves changing something in an existing environment to encourage more natural behaviour. Examples of this can be seen in most European/American zoos, where man-made termite mounds have been built into chimpanzee enclosures to encourage more natural foraging activities and tool-use[8]. This approach has also been extended to farmed animals. Pigs have been provided with chains to manipulate, straw to root and balls to 'play' with. The addition of such objects has been claimed to reduce excitability, making the pigs easier to move and manage. It is claimed in addition that 'toys' help reduce the stressfulness of the often barren environment in which pigs live[5].

Modification of the environment can help solve the problems associated with some very intensive farming conditions. For example, in Europe and America over 90% of laying hens are housed in 'battery cages'[9]. These are so called because a large number of cages are placed in tiers within one hen-house. Layers are kept in groups of 3–6, with an individual space allowance of 300–450 cm^2 per bird. The cages pose problems for the birds since, although they are provided with adequate food and water, they are not able to fly, perch, walk about, peck the substrate, nest, or escape from each other.

There has been a widespread call for a ban on these cages but alternative forms of managing large flocks of birds have not yet been fully developed. The Edinburgh modified cage is an attempt to improve the hens' welfare by enabling them to live in larger cages, fitted with perches, nest boxes and dust baths. Production is above average in these cages, and all facilities are heavily used. Feather, foot and claw damage is also reduced. It is felt that further development of such modified cages would lead to vast improvements in hen welfare.

In conclusion, environmental design is a broad subject encompassing the physical, social as well as developmental needs of the animal. Designing animal environments to take these needs into consideration, as well as the needs of the humans operating the system, is a difficult task, and in order to do it effectively we must have an understanding of the impact the environment can have on the well-being of the animals. A logical sequence of events leading to a new system would, therefore, be: (i) define the behavioural requirements; (ii) determine if existing systems contravene these requirements and, if so, determine the consequences for the animal; and (iii) if there is a need following (ii), then a new system(s) requires to be designed taking into account behavioural requirements and other management and economic considerations.

1. Wiepkema, P. R., Broom, D. M., Duncan, I. J. H. and Van Putten, G., Abnormal behaviour in farm animals. A report of the commission of the European communities, EEC Publications, Brussels, 1983, p. 16.
2. Mason, G. J., *Anim. Behav.*, 1991, **41**, 1015–1037.
3. Stolba, A. and Wood-Gush, D. G. M., *Ann. Rech. Vet.*, **15**, 287–298.
4. Appleby, M. C. and Smith, S. F., *Br. Poultry Sci.*, 1991, **32**, 667–678.
5. Grandin, T., *Hog Farm Management*, pp. 6–8.
6. Waran, N. K. and Broom, D. H., *Anim. Prod.*, 1991, **56**, 115–119.
7. Dawkins, M., *Behav. Brain Sci.*, **13**, 1–61.
8. Shepherdson, D., in *Creative Conservation: Interactive Management of Wild and Captive Animals* (eds Mace, G., Olney, P. and Feistner, A.), Chapman & Hall, London, 1993.
9. Appleby, M. F., Hughes, B. O. and Elson, H. A., *Poultry Production Systems: Behaviour, Management and Welfare*, C. A. B. International, Wallingford, Oxon, 1992.

Stereotypies in the stabled horse: Causes, treatments and prevention

L. Winskill, N. K. Waran, C. Channing and R. Young

IERM, University of Edinburgh, Kings Buildings, West Mains Road, Edinburgh, EH9 3JG, UK

The environment for most stabled horses is far removed from that of its wild ancestors. The horse is a social animal, and in natural groups in the wild it lives in harems, usually consisting of one stallion, a few mares and all their offspring up to two-years old[1]. They restrict their movements to a specific home range within which the major requirements are food, water and shelter. Home ranges of free-ranging horses vary greatly in size and are correlated with resource availability[2]. In contrast, the stabled animal is removed from this normal open range and often deprived of social contact. Many performance horses (e.g. racehorses, show jumpers, etc.) in Europe and also in India are confined for up to 24 h a day with the exception of short periods of time reserved for exercise, and others are turned out to pasture for merely a few hours a day.

This restricted environment often has a significant effect on the normal time budget of the horse as it can no longer spend the majority of the day foraging and expending energy looking for suitable resources[2]. Feeding with concentrated food also has consequences for the horse, since it eliminates the additional time that confined animals can occupy themselves with ingestion and manipulation of food[2]. Free-ranging horses spend a high proportion of their time grazing. A study by Boy and Duncan[3] on free-ranging feral Camargue ponies during spring and summer revealed that 57.5% of their time was spent grazing. This proportion of time is likely to be greater in the winter months, when grass is scarcer. In the stabled environment, however, if fed on hay, horses will spend 40% of their day eating, which is reduced by one tenth if fed chiefly on concentrates[4].

As already highlighted, the stabled horse is often deprived of social contact, both visual and physical. The design of the majority of housing for horses does not take into consideration their behavioural needs, often enclosing the animal in small individual loose boxes or stalls, where it is deprived of social contact and exercise. Rees in her book *The Horse's Mind*[1] describes the confined horse as being like 'a fish out of water'. She notes that many learn to accept confinement, but others kept in these unnatural conditions, deprived of freedom,

exercise and companions, become excitable and irritable. It is in such conditions that behavioural problems may arise.

This article discusses the incidence of behavioural problems of the stabled horse giving examples of those most commonly found and how they might be treated and prevented from occurring. More general factors relating to the occurrence of these problems are also mentioned.

Behavioural problems and their incidence

Stabled horses can be seen to exhibit a variety of behavioural problems or 'stereotypies' such as crib-biting, weaving, wind-sucking and wood-chewing. Such behavioural anomalies are often wrongly referred to as 'vices', a term which implies that the blame should be attributed to the horse itself and not to the management procedures adopted or the housing imposed by humans[5,6]. An example of such misuse of terminology is highlighted in the paper by Krzak *et al.*[7] regarding wood-chewing of stabled horses, in which they state that: 'The owner of a wood-chewing animal must absorb the costs of replacing damaged fences and stalls and of preventative measures such as covering wood surfaces with metal, wire or taste repellents. For these reasons, horses that exhibit this 'vice' are of reduced value and may not be saleable'. The welfare of the animal is not mentioned; merely the cost and inconvenience brought upon the owner are highlighted. Fraser and Broom[5] believe that such attitudes have been an important factor in the perpetuation of many systems which result in poor welfare, while Luescher *et al.*[8] suggest that the term 'vice' should be reserved for nuisance behaviours related to temperament such as biting, kicking and rearing.

Some behavioural problems of the stabled horse are more accurately referred to as stereotypies, e.g. 'repetitive, invariant behaviour with no obvious goal or function'[9], and are indicative of a situation in which the animal lacks a certain degree of control over its environment[5]. Stall-walking, crib-biting, wind-sucking, weaving, swaying, pawing, stall-kicking, head-shaking, head-nodding, crib-whetting and tongue-dragging are behaviours that have received detailed description, and are examples of movement sequences which horses repeat with very little variation in form[6]. Stereotypies have been recognized and much researched in many confined animal species, for example, zoo animals[10], pigs[11], sheep[12] and domestic fowl[13], but as emphasized by Rees[1], such behaviours in stabled horses seldom receive the same attention.

The proportion of horses exhibiting stereotypies varies among studies, which makes it difficult to assess the extent of the problem. A study by Kennedy *et al.*[14] revealed that 18.4% of the 49 horses observed at two riding schools and four private livery yards showed clearly recognizable stereotypies. This figure was even thought to be an underestimate, as the staff at the stables described more horses as stereotypy performers than those identified during observation. Kiley-Worthington and Wood-Gush[15] stated that in some of the racehorse stables which they visited, all the horses were exhibiting stereotypies, and in another stable 26% of the horses were seen to perform a stereotypy within only half an hour of observation. A recent article[8] quoted that over 15% of domesticated horses exhibit stereotypies.

These figures, however, should all be questioned due to the small sample sizes involved in the studies. In an investigation involving 1035 thoroughbred horses aged 3–8 years in Italy, Vecchiotti and Galanti[16] report the incidence of cribbing, weaving and stall-walking as 2.4%, 2.5% and 2.5%, respectively, of all behaviours. Similarly, a recent study carried out on 21 thoroughbred horses in the north and centre of Italy showed that 67 out of 1180 (5.7%) horses showed a behavioural problem[17]. These lower percentages are perhaps a better representation, as a larger number of horses were analysed. However, until a study is undertaken comparing the incidence of stereotypies between differing breeds, ages and sexes of horses with varying types of yard (e.g. breeding, racing, livery or riding school), it will remain difficult to quote an accurate figure representing the extent of occurrence of stereotypies in the stabled horse. Another factor which will have an effect on studies involving survey questionnaires distributed to horse owners is a reluctance on the part of some owners to admit that their horse is a stereotypic performer (as this often decreases the value of the animal), leading to underestimation of the affected numbers.

Stereotypies: examples, causes and treatments

Quantitative evidence relating to the causal factors of stereotypies in stabled horses is very scarce in the scientific literature, the majority being purely anecdotal in nature. With the exception of more recent studies[17,18], research in this area has involved either small sample sizes[19] or has consisted of a collection of individual case studies[14]. This lack of evidence implies that the causes of abnormal behaviour are still unclear, and more research is still required.

Wood-chewing

Wood-chewing is a normal component of a free-ranging horse's feeding behaviour, especially when preferred grasses are less abundant during late autumn and winter[20]. Wood-chewing is, however, a common problem in the stabled environment, where a horse may chew 0.5 kg of wood a day[5]. Coupled with the financial burden on

the owner to replace fences and stable boards, wood-chewing can also cause damage to the animal itself, largely by causing excessive wearing of teeth[21]. It can pose more serious consequences such as gastrointestinal disease[22], and small splinters of wood or nails may puncture the oesophagus or mouth[5,21], and may also lead to the formation of phytobezoars or enteroliths[21].

The association of wood-chewing with pelleted feeds has been well documented[22]. The exact cause of wood-chewing associated with these feeds is unknown although reduced cecal pH and relative lack of fibre or 'chewing time' or both have been suggested[5,22]. The horses' natural diet consists of an almost continuous consumption of roughage, a large amount of which is required by the system for proper digestive fermentation. Willard *et al.*[19] undertook an experiment to determine the effect of diet on cecal pH and feeding behaviour of horses. For the study, three fistulated horses were used, and their dietary treatments were hay, concentrate and concentrate plus cecal infusions of Na_2CO_3. Horses fed on hay were found to spend significantly more time eating feed and significantly less time chewing wood, in coprophagy, in searching for feed and in standing than those receiving the concentrate diet. The amount of wood chewed per day was significantly greater for horses fed on concentrate diet and the Na_2CO_3 infusions appeared to reduce the amount of time spent chewing wood by the concentrate-fed horses, but the difference was not significant. They state that their results may indicate that increased cecal acidity and a narrowed acetate–propionate ratio, as was noted with the all-concentrate diet, influences the horse's desire to chew wood and practice coprohagy.

Doubts, however, have to be raised about the significance of these results, firstly due to the small sample size used (three horses), and secondly as only two feeding regimes were used: 8 kg of mixed grass legume hay and 6 kg of a sweet feed mixture fed at 12 h intervals. A combination of hay and concentrate may have been more appropriate to normal feeding practices. Another factor is that the authors do not mention whether these quantities of hay and concentrate represent the same percentage of body weight, which will have an effect on the results.

The results of this experiment contrast with those of a study carried out by Jackson *et al.*[23], where horses were subjected to one of the three treatments. In the first, the animals were fed long-stemmed alfalfa hay twice a day; in the second they were fed alfalfa cubes twice a day, and in the third they were fed alfalfa cubes three times a day. Both the cubes and hay were fed at 2.5% of body weight per horse per day. It was shown that there was no difference in the amount of wood chewed, despite the fact that animals in treatment 2 spent less time eating than those in treatments 1 and 3, and the cubed alfalfa would have contained less roughage than the long-stemmed hay (C. Morgan, personal communication). The difference between the two studies could possibly be explained by the fact that horses in the latter experiment were able to socialize, whereas those in the former were maintained in metal stalls, where boredom could also be a factor predisposing the horses to chew wood. Furthermore, reduced cecal pH and lack of fibre or chewing time[19, 22] suggest that the degree of intestinal fill provided by the diet should also be investigated.

Exercise of the stabled horses is also believed to reduce wood-chewing by reducing boredom[7]. A study undertaken by Krzak *et al.*[7] demonstrated that horses chewed more wood when they did not receive exercise than when they received 45 min of exercise on a mechanical treadmill followed by a further 45 min in a paddock with bare soil in either the morning or the afternoon. The greatest amount of wood-chewing occurred during the late night and the early morning when distractions were minimal. The authors believe that this provides some support for the boredom theory, and also suggest that hunger may have been associated with wood-chewing in this particular study, as its incidence was seen to increase several hours after the afternoon feed. Feed restriction has already been found to be related to the performance of oral chain manipulation in tethered gilts[11], stereotyped pecking at nonfood objects in broiler breeders[24] and abnormal behaviours, including chewing wood in individually penned sheep[12]. Frustration of feeding motivation has been suggested as an underlying cause of oral stereotypies in pigs[25]; however, in this particular study[7], wood-chewing did not appear to be related to the time of feeding, as oral contact was not associated with the afternoon feeding and did not increase immediately before or after the morning feeding.

Finally, weather has been associated with the amount of wood-chewing. In the study by Jackson *et al.*[23], the incidence of wood-chewing was seen to increase, in both wet and cold weather.

The most common method employed to prevent wood-chewing is to cover all wood surfaces with a noxious substance such as creosote[5], carbolium or lemon-flavoured oils[22]. Such practices, however, only prevent the performance of the behaviour and fail to identify the underlying cause. It is obviously important that the animal receives a correct diet in terms of protein and energy, and should contain more than 50% roughage, or there will be a risk of failing to satisfy its needs for fibre fill and chew time[3]. If it is essential that the horse be confined, adequate exercise should be given, and attempts made to relieve boredom; for example, by providing a companion animal such as a goat[26]. Providing a sheep as a social companion proved successful in an experiment looking at social facilitation in pigs, in which one of the treatments involved housing the animals in isolation[27]. The inclusion of toys has also been suggested[26]. Since wood is a natural component of a

horse's diet, the problem cannot be entirely eliminated[26].

Crib-biting and wind-sucking

Crib-biting and wind-sucking are probably the most easily recognizable sterotypies in the stabled horse[14]. Wind-sucking involves opening the mouth, contracting the pharyngeal muscles, flexing the neck muscles and swallowing air, usually making a characteristic grunting noise. Crib-biting (or cribbing) is characterized by the horse resting its upper incisors on a solid object such as a door or a manger. It then leans back, arches its neck and engulfs a bolus of air (aerophagia)[28,29]. In rare cases, the mouth may be placed against the knees or cannons[5]. Cribbing is usually associated with stabled horses; however, some will crib when turned out to pasture[28]. Both crib-biting and wind-sucking can be detrimental to the health of the animal. Crib-biting can cause excessive wear of the incisor teeth, in some instances they may be worn down to the gum line[26,27], and both crib-biting and wind-sucking can cause the musculature of the throat to increase in size, and make the animal more susceptible to flatulent colic[5,29]. Horses which practise aerophagia intensely may reduce their intake of feed, which will ultimately lead to a deterioration of their physical condition[5].

Heredity has been thought to play a role in the frequency of stereotypies. A survey was undertaken by Vecchiotti and Galanti[16] in which 80 trainers were interviewed on the occurrence of anomalous behaviours, involving a total of 1035 thoroughbred horses. The incidence of crib-biting (also weaving and stall-walking) in related horses was found to be higher than that in the surveyed population. The authors suggested that this was due to the involvement of inheritance in each of the three behaviours. However, as discussed by Broom and Kennedy[6] this could simply be due to the fact that many related horses, particularly thoroughbreds, are kept in similar housing conditions which may give rise to the performance of stereotypies.

In other species, it has been shown that stereotypies are associated with opioid activity, which may have the effect of quietening the performer[30]. When naloxone is administered to stereotyping sows, the behaviour is stopped[5]; narcotic antagonists were also found to prevent crib-biting in horses[29]. Baseline β-endorphin levels in cribbing horses were found to be half of those in non-cribbing controls. This was thought to be due to cribbing horses having an impaired release of β-endorphin, which in turn increases the sensitivity of opioid receptors (Gillam *et al.*, 1994, unpublished).

In a study investigating the effects of different diets on the frequency of cribbing behaviour, feeding grain or sweetened grain rations was seen to cause a significant increase in cribbing frequency (Gillam *et al.*, 1994, unpublished). This is thought to be due to the palatability of these particular rations, since palatable food has been shown to release β-endorphins and enhance stereotypic behaviour patterns in the horse[29]. Kennedy *et al.*[14] undertook a study in an attempt to discover factors relating to both management, and horses' characters which may be linked to the development of stereotypies. Many of the individual animals involved who performed crib-biting were seen to exhibit this behaviour around the time of feeding. A proportion was seen to exhibit crib-biting post-feeding, which could possibly be related to the palatability of the food[29]. Heifers tied up in stalls have also been found to spend most of their time performing stereotypies 2–4 h after feeding[31]. Other bouts of crib-biting in this study, however, were linked to the anticipation of food prior to feeding[14]. Pigs exposed to an intermittent delivery of food with a chain in front of them were seen to develop stereotypic chain manipulation in the interval between food deliveries, and displayed a lower plasma cortisol level than in the absence of the chain, suggesting a stress-reducing role of chain manipulation when the pigs were anticipating the delivery of food[32], which could possibly be an explanation for the increase in the frequency of crib-biting seen in horses prior to feeding time.

Feeding has been associated with the performance of oral stereotypies in a number of other species; for example, wire-chewing in commercially housed rabbits and bar-biting in heifers[31], sheep[33] and pigs[11]. When on high feed intakes, tethered gilts performed a lower incidence of repetitive behaviours such as bar-biting and chain-chewing, which increased when fed a small ration[11]. Bar-biting, crib-biting and other stereotypies are more frequent in the absence of straw or other manipulable material[34], so it appears that the type and bulk of forage offered as well as the ability to manipulate are important[5].

Traditionally, the control of crib-biting and wind-sucking has been directed towards preventing the behaviour by mechanical or surgical techniques[26]. The cribbing strap is the most commonly used mechanical control device. This consists of a 4 cm leather strap fastened around the throat, tightly enough to make arching of the neck uncomfortable, but not so tight as to interfere with respiration[5,28]. In some designs, there is a metal 'gullet-piece' which has a recess into which the windpipe fits, or a pressure plate containing one to three sharp spikes[5,28]. These straps can be buckled so tightly that they interfere with normal eating, and in one noted case a tightly applied strap was associated with acute chondritis of the larnyx[28]. Another preventive measure consists of a hollow cylindrical perforated bit, which prevents the animal from making its mouth airtight[5].

Surgery has been performed in both wind-sucking and crib-biting horses[5,28,35,36]. It is, however, a very drastic

measure to undertake in order to prevent behaviours caused largely by management practices imposed by man. The success rate of operations is also fairly low, and the majority of animals responding to these methods of prevention are those which have not practised the anomaly for long periods[5].

Attempts such as surgery or the use of a cribbing strap when used to prevent the performance of a behaviour prevent only its manifestation without eliminating the causal factors involved. Such measures may cause additional frustration for the animal, which may perform another stereotypic behaviour that it is not restricted from expressing, further compromising its welfare. As recommended for animals performing wood-chewing, the housing and management of the horse needs to be improved. A simple task such as feeding at random times throughout the day may help in reducing the expression of crib-biting and wind-sucking which occur prior to feeding. However, in their most severe forms, wind-sucking and crib-biting may continue to be performed when turned out to pasture (personal observation), emphasizing the persistence of stereotypies once developed[9]. In such circumstances where the behaviour continues despite recommended changes in feeding and other management practices, such as the provision of social companionship and increased exercise regimes, the use of narcotic antagonist drugs such as naloxone and naltrexone, may play a role[29], although this could be argued by some to be no better than surgery or physical restriction.

Weaving and stall-walking

Weaving and stall-walking have been described as locomotor activities which have become fixations[26]. The main feature of weaving is that the horse swings the head and neck from side to side, whilst shifting its weight from one foreleg to the other[5,22]. This is similar to the rocking behaviour shown by monkeys in captivity[5] and in institutionalised humans[9]. Head-swaying has also been observed in singly caged rabbits, where its function has been suggested as a mechanism to relieve boredom[37]. Stall-walking describes a form of stereotypic pacing, in which the horse constantly walks around the perimeter of the stable. It is thought to resemble weaving as both behaviours involve accurate precision and repetition[5]. Both activities can lead to weight loss, but particularly 'stall-walking', where vast quantities of energy can be depleted[5,26]. Stall-walking can also cause flexion of the spine, especially if the area available to the animal is particularly small[5]. Apart from its aesthetic appearance, the performance of weaving is disliked by horse owners as it can cause uneven musculature development in the neck, which can have an effect upon performance.

There is little information available regarding the causal factors of these behaviours, although frustration, boredom due to lack of variety in the environment and lack of exercise have been suggested[5,26]. Lack of variety in the environment has been found to be a contributing factor in the occurrence of stereotypic pacing in captive-felids, as the frequency of this behaviour in captive leopards was found to decrease significantly after enrichment of the environment[10]. In many captive carnivore species, stereotypic pacing is thought to be caused by the build-up of frustration due to being unable to perform predatory behaviour[10]. Stereotypic pacing in captive American black bears was seen to reduce dramatically when food was hidden in manipulative objects[38], and in captive polar bears expectation of food was found to influence the frequency of pacing behaviour[39]. Feeding could also be one of the contributing factors influencing the onset of stall-walking in the stabled horse, but there is no evidence to suggest this, more research obviously being required in this area.

There are not many measures undertaken to prevent these behaviours. A metal grate is often attached to the stable door to prevent horses from weaving, but the most obvious method of preventing an animal from developing these stereotypies would be to turn them out to pasture, as they appear to be linked to the need to exhibit locomotory behaviour.

General factors relating to the occurrence of stereotypic behaviours

As previously mentioned in relation to crib-biting and wind-sucking, inheritance is thought to play a role in the performance of stereotypies in horses. There is, however, little evidence to accompany this hypothesis with the exception of the survey of thoroughbred horses in Italy by Vecchiotti and Galanti[16], in which they revealed a genetic component in the causation of crib-biting, weaving and stall-walking. A study in Japan[6] came to a similar conclusion. This is certainly apparent in dogs, in which particular breeds seem to be predisposed to abnormal behaviours (e.g. whirling in Bull Terriers)[8]. Kiley-Worthington[40] also notes that stereotypies seem to be more prevalent in thoroughbred and Arab-type horses, which suggests a genetic factor, although she suggests that this could be due to the restricted environments that these 'warm-blooded' breeds are kept in.

It has been proposed by a number of authors that stereotypies can be learnt by imitating neighbouring stereotypic performing animals, particularly between mothers and their offspring[5,8,40]. Stereotypic behaviour in bank voles was more rapidly developed in those individuals which were visually exposed to a demonstrator already performing stereotypies[41]; similarly, neighbour-

ing stereotyping sows were seen to influence the amount of stereotypic behaviour performed by tethered gilts[42]. However, in an experiment designed to determine whether a horse could learn the position of grain in one or two buckets by watching another horse, Baker and Crawford[43] showed that no observational learning was demonstrated by the animals. The reason why neighbouring horses develop stereotypies is, therefore, more likely to be as a result of them being kept in the same inadequate environment rather than due to their ability to copy.

Abrupt weaning has also been suggested as a possible factor relating to the onset of the performance of stereotypies, as a survey of thoroughbred breeding farms in Italy showed that 23.8% of horses less than two years old showed abnormal behaviour[17]. These behaviours were seen to commence immediately after the post-weaning period, and as the animals were turned out until the training period, the authors believe that the abrupt weaning could be a stressor for the foals, which could in turn lead to behavioural problems.

The management of young horses in training was the major factor linked to an increase in the prevalence of abnormal behaviour in a survey sampling the largest uniform population of thoroughbred horses yet to be studied, in relation to the performance of weaving, box-walking, wind-sucking and crib-biting[18]. The same study found that the risk of horses performing abnormal behaviour increased when the total number of horses on the yard was less than 75, in box designs which minimized social contact between horses, when bedding types other than straw were used, and when the amount of feed was less than 8.8 kg per day.

Conclusion

In order to prevent the development of stereotypies in the stabled horse, an appreciation of the type of social and physical environment the horse has evolved to live in is obviously required[40]. This knowledge should also be taken into consideration in future housing design; for example, by providing more space and social companions in the form of group housing. It will, however, probably prove difficult to persuade owners to take into consideration their animal's behavioural requirements as horses are very valuable animals and are too often 'wrapped in cotton wool'. This is highlighted in a survey of thoroughbred breeding farms in the north and centre of Italy, where stallions were found to perform a higher frequency of abnormal behaviours than mares[17]. The authors suggest that the reason for this is the different management of the two sexes. This is probably true, since stallions are worth vast sums of money and are nearly always kept in single stables, with no access to pasture for fear of theft or injury. Horse owners, therefore, need to be educated about the psychological causes of stereotypies, as most are treated only because they are not aesthetically pleasing and may decrease the value of the animal.

If horses are to be confined, then every effort must be made to prevent the occurrence of these abnormal behaviours using suggestions outlined in the article; for example, by providing hay *ad libitum*, providing a social companion and adequate exercise. There is also scope for the provision of enrichment devices such as the 'Edinburgh foodball', which has already proved to be a successful foraging device for pigs[44].

1. Rees, L., *The Horse's Mind*, Stanley Paul, London, 1993.
2. Waring, G., *Horse Behaviour*, Noyes Publications, New Jersey, 1983.
3. Fraser, A. F., *The Behaviour of the Horse*, CAB International, Wallingford, 1992.
4. Boy, V. and Duncan, P., *Behaviour*, 1979, **71**, 187–202.
5. Fraser, A. F. and Broom, D. M., *Farm Animal Behaviour and Welfare*, Bailliere, Tindell, 1990.
6. Broom, D. M. and Kennedy, M. J., *Equine Vet. Educ.*, 1993, **5**, 151–154.
7. Krzak, W. E., Gonyau, H. W. and Lawrence, L. M., *J. Anim. Sci.*, 1991, **69**, 1053–1058.
8. Luescher, U. A., McKeown, D. B. and Halip, J., *Vet. Med.*, 1991, **May**, 527–530.
9. Mason, G. J., *Anim. Behav.*, 1991, **41**, 1015–1037.
10. Carlstead, K., Brown, J. L. and Seidensticker, J., *Zoo Biol.*, 1993, **10**, 3–16.
11. Appleby, M. C. and Lawrence, A. B., *Anim. Prod.*, **45**, 1987, 103–111.
12. Marsden, D. and Wood-Gush, D. G. M., *Anim. Prod.*, 1986, **42**, 157–159.
13. Hughes, B. O., *Appl. Anim. Ethol.*, 1983, **11**, 45–53.
14. Kennedy, M. J., Schwabe, A. E. and Broom, D. M., *Equine Vet. Educ.*, 1993, **5(3)**, 142–147.
15. Kiley-Worthington, M. and Wood-Gush, D. G. M., in *Behavioural Problems* (ed. Houpt, K. A.); *Current Therapy in Equine Medicine* (ed. Robinson, N. E.), W. B. Saunders Company, Philadelphia, pp. 131–134.
16. Vecchiotti, G. G. and Galanti, R., *Livest. Prod. Sci.*, 1986, **14**, 91–95.
17. Canali, E. and Borroni, A., in Proceedings of the International Congress on Applied Ethology, Humbol University, Berlin, 26–30 July, 1993; *Appl. Anim. Behav. Sci.*, 1994, **40**, 74.
18. McGreevy, P. D., Cripps, P. J., French, N. P., Green, L. E. and Nicol, C. J., in Proceedings of the 28th International Congress of the International Society of Applied Ethology, Research Centre, Foulum, Denmark, 3–6 August, 1994.
19. Willard, J. G., Willard, J. C., Wolfran, S. A. and Baker, J. P., *J. Anim. Sci.*, 1977, **45**, 87–92.
20. Tyler, S. J., *Anim. Behav. Monogr.*, 1972, **5**, 85–196.
21. Hintz, H. F., *Equine Pract.*, 1992, **14(3)**, 6.
22. Ralston, S. L., in *Behavioural Problems* (ed. Houpt, K. A.), and *Current Therapy in Equine Medicine* (ed. Robinson, N. E.), W. B. Saunders Company, Philadelphia, pp. 123–126.
23. Jackson, S. A., Rich, V. A., Ralston, S. L. and Anderson, E. W., *J. Anim. Sci.*, 1985, **59** (Suppl. 1), 152.
24. Kostal, L., Savory, C. J. and Hughes, B. O., *Appl. Anim. Behav. Sci.*, 1992, **32**, 361–374.
25. Rushen, J. P., *Appl. Anim. Behav. Sci.*, 1985, **14**, 137–147.
26. Ralston, S. L., *Compend. Continu. Educ.*, 1982, **4**, 5152–5159.
27. Hsia, L. C. and Wood-Gush, D. G. M., *Anim. Prod.*, 1983, **37**, 149–152.
28. Monin, T., *Compend. Continu. Educ.*, 1982, **4**, 569–572.

29. Dodman, N. H., Shuster, L., Court, M. H. and Dixon, R., *Am. J. Vet. Res.*, 1987, **48**, 311–319.
30. Wiepkema, P. R. and Koolhaas, J. M., *Stress Anim. Welfare*, 1993, **2**, 195–218.
31. Redbo, I., *Appl. Anim. Behav. Sci.*, 1990, **26**, 57–67.
32. Dantzer, R. and Mittleman, G., in *Stereotypic Animal Behaviour: Fundamentals and Applications to Welfare* (eds Lawrence, A. B. and Rushen, J.), CAB International, Wallingford, pp. 147–171.
33. Stauffacher, M., *Anim. Welfare*, 1992, **1**, 105–125.
34. Fraser, D., *Anim. Prod.*, 1975, **21**, 59–68.
35. Firth, E. C., *Vet. Rec.*, 1980, **106**, 30–32.
36. Frauenfelder, H., *Equine Vet. J.*, 1981, **13(1)**, 62–63.
37. Morton, D. B., Jennings, M., Batchelor, G. R., Bell, D., Birke, L., Davies, K., *et al.*, *Lab. Anim.*, 1993, **27**, 301–329.
38. Carlstead, K., Seidensticker, J. and Baldwin, R., *Zoo Biol.*, 1991, **10**, 3–16.
39. Ames, A., *Biologist*, 1993, **40(3)**.
40. Kiley-Worthington, M., *Equine Pract.*, 1983, **5**, 34–40.
41. Cooper, J. J. and Nicol, C. J., in Proceedings of the UK Regional Meeting of the International Society for Applied Ethology, Royal Veterinary College, London, 2nd December, 1992; *Appl. Anim. Behav. Sci.*, 1993, **38**, 80.
42. Appleby, M. C., Lawrence, A. B. and Illius, A. W., *Appl. Anim. Behav. Sci.*, 1989, **24**, 147–155.
43. Baker, A. E. M. and Crawford, B. H., *Appl. Anim. Behav. Sci.*, 1986, **15**, 7–13.
44. Young, R. J., Carruthers, J. and Lawrence, A. B., *Appl. Anim. Behav. Sci.*, 1994, **39**, 237–247.

ICI India Limited

ICI India Ltd. is one of the leading chemical companies engaged in diverse businesses including Pharmaceuticals, Paints, Explosives, Rubber Chemicals and Surfactants. The company is presently looking for:

SCIENTISTS

for its prestigious R&D Centre located at Thane near Bombay. The Centre has excellent facilities and high calibre staff, and undertakes national and international projects. Applicants should have a Ph.D in SYNTHETIC ORGANIC CHEMISTRY with relevant postdoctoral experience. Exposure to Industrial research would be an advantage, but is not essential. Only persons with proven track record of excellence should apply.

The company offers professionally challenging and stimulating work environment, performance-based career growth & an excellent remuneration package including company leased housing.

Please apply within 10 days to: The Personnel Manager, ICI India Ltd., P.O. Box 87, Thane-Belapur Road, Thane - 400 601, Maharashtra: Fax: (022) 763 2889.

Seamless chemical engineering science: The emerging paradigm*

R. A. Mashelkar

I consider it to be a matter of great honour that I have been asked to deliver the ninth P. V. Danckwerts Memorial Lecture. Danckwerts' contributions to the evolution of Chemical Engineering Science (CES) are monumental. He set the pace and the tone and defined the features of CES in the early fifties through his path-breaking and incisive contributions. His work has continued to inspire generations of chemical engineering scientists (CE scientists). I certainly happen to be one of them. It is, therefore, a special privilege that I have this opportunity to pay my own tribute to Danckwerts. Today I wish to speak to you on an emerging paradigm – the seamless chemical engineering science.

How did Danckwerts himself view CES? I searched and found a statement in his presidential address delivered to the Institution of Chemical Engineers almost three decades ago. He said: 'I shall use the phrase Chemical Engineering Science to mean the aspects of science developed by chemical engineers for their own purposes in fields not covered by other branches of technology'[1]. He also went on to alert us: 'It will be a great mistake to think of the content of chemical engineering science as permanently fixed. It is likely to alter greatly over the years in response to the changing requirements of industry and to new scientific discoveries and ideas for their application'. He was absolutely right. I wish to refer to those explosive advances in science which are reshaping our lives. I also wish to emphasize how these advances in adjacent sciences will force us to reconceptualize CES of tomorrow. I particularly want to focus on the merging boundaries between CES and the adjacent sciences and also on the need to build new bridges with our neighbours.

We look around and see that discrete boundaries no longer exist between various natural sciences such as physics, chemistry, biology, mathematics, etc. Let us remember that the 1985 Nobel Prize for chemistry was actually awarded to two mathematicians! Take even subdisciplines such as inorganic chemistry, biological chemistry, organic chemistry and physical chemistry. The conceptual boundaries between these disciplines have already vanished. The Nobel Prize winning work of Lehn[2] was a testimony to this realization.

Is only science going borderless? Of course not. The borderless world has already dawned on us. Borderless organizations are here to stay too. I wish to quote from the speech by Jack Welch, the Chairman and Chief Executive Officer of General Electric (GE), whom we all respect for his dynamism and vision. In his lecture on 'Global Competitiveness' delivered last month, he talked about GE's strong commitment to the boundaryless behaviour and said, 'Boundaryless is the language, the behaviour definer, the culture, the soul of a true global enterprise. It ignores geography, borders, accents, currencies, and unites people of all cultures'. When asked about what happens to those GE people who did not accept this boundaryless behaviour, he responded by saying, '(we) tell them they have to change, or else if they don't change, they have to part company with us. And we do have them part company – lots of them over the question of boundaryless behaviour, of being open'. Strong words indeed – but it shows the way it works in industry. Why does it do so? Let me answer this.

Successful commercialization of modern scientific discovery today demands cross-disciplinary collaboration between scientists and engineers from different disciplines. Advances are generally the sum total of numerous creative ideas and interdisciplinary cooperation. In industry, individuals with diverse scientific and technical backgrounds tend to work together in a mission mode. The driving force for industry is continuous innovation to displace one's own product by oneself – with the well-known fear of someone else displacing it otherwise for them. The concept of interdisciplinary research and development has not gained as much ground in academia as one would have wished. And this is really sad because innovation is not a unidimensional process. It is comparable to the intermeshing gears of a clock. The challenge before us is to make this intermeshing happen.

Why did I decide to speak on this topic? I find myself in an interesting situation. I head the National Chemical Laboratory (NCL) in India. I have to deal with multidisciplinary teams of over 500 research scientists and engineers. I find myself communicating with physical

R. A. Mashelkar is in Council of Scientific and Industrial Research, New Delhi 110 001, India.

*Ninth P. V. Danckwerts Memorial Lecture presented at Glazier's Hall, London, UK, 9 June 1994.

Reprinted from *Chemical Engineering Science*, Vol. 50, R. A. Mashelkar, 'Seamless Chemical Engineering Science: The emerging paradigm,' page 1–22, 1994, with kind permission from Elsevier Science Ltd., The Boulevard, Langford Lane, Kidlington OX5 1GB, UK.

chemists, organic chemists, inorganic chemists, etc. I deal with molecular biologists, biochemists, bio-organic chemists too. I also deal with materials scientists, applied mathematicians, NMR experts and so on. Our CES group is placed amidst all these groups and it has to interact with all these disciplines. This interaction is crucial to meet NCL's charter, which is to develop globally competitive products, processes and technologies. Interestingly, we have found that we succeed in this only when we come together as a team, leaving aside all the conceptual and cultural barriers between different disciplines. But I must also honestly admit that I have found it to be a hard act to put together. I thought I would share the agony and ecstasy that I have experienced during all these years of going 'boundaryless' with you.

But before I begin to address the principal issues, I must put the role of an engineering scientist, as I see it, in perspective. I would like to give an analogy with the medical school. We have there basic scientists, who generate new knowledge, but it is knowledge for its own sake. Then we have the hard-core practising clinicians, who use the established tools in their everyday practice. But in between these two communities, there are those academic clinicians. They are the ones who understand the language of the basic scientists and they are the ones who know the demands of the profession, as reflected by practising clinicians. What these academic clinicians do is to use the methods of science to develop new tools and new paradigms that the practitioners would find most helpful. They also communicate the needs of these practitioners to the scientists. The academic clinicians thus play this great role of facilitators. To my mind, a CE scientist is like an academic clinician, a facilitator. His task is to look at the continuous advances that take place in basic sciences such as chemistry, biology, physics, etc. He would harness this knowledge to develop new tools, which hard-core chemical engineers would put to use. The richness of our chemical engineering profession, therefore, depends very critically on the vitality of the CE scientists and also on the rich variety of skills, tools, etc., that they bring into the profession.

I will give my personal perceptions about the following issues. What is the emerging relationship between CES and adjacent disciplines? What driving force is shaping this relationship? How do we respond to the explosion of knowledge in adjacent sciences? What will be the new role of a CE scientist *vis-à-vis* chemical engineering and other disciplines of engineering and sciences? What new tools and techniques can we adopt from other disciplines and, in turn, what could we give to other disciplines? Danckwerts practically defines Chemical Engineering Science as 'applying methods of science to distinctive chemical engineering problems'. Is this definition valid anymore? What should be the new definition?

A word about how I have organized my lecture may be in order. This is certainly not a lecture on 'new frontiers in CES', since these frontiers have been widely discussed and elaborately defined by the Amundson report[3]. I am also not getting involved in a discussion of the 'core' vs 'frontiers' issue, which has again been dealt with by some of our very best scholars and thinkers[4,5]. I am also not speaking about the future of fields such as mass transfer, reaction engineering, applied thermodynamics and so on. Furthermore, although I want to deal with the issue of merging boundaries with other disciplines, I am not going to create distinct boxes such as CES–physics interface, CES–biology interface, etc. To me, the boundaries between other disciplines have already vanished. I will, therefore, deal with only some 'generic issues' that should be of concern to us.

This lecture is admittedly a narrow view from a window. I also have the inescapable personal bias, which might show through the examples and the illustrations that I have chosen to give, and also the recommendations that I have dared to make despite my limitations. However, I thought it is worth putting these thoughts on record, since this is the time to do it. Tomorrow might be too late!

Nonlinear dynamics – New interdisciplinary challenges

Complexity and nonlinearity reside in much that a chemical engineer encounters. His ability to acquire new tools and knowledge by exploiting the contemporary advances in physics has undergone a sea change in recent years. A wide variety of features ranging from steady state to multiple steady states to oscillations to chaotic dynamics and spatiotemporal patterns[6,7] have become more commonplace observations. The advent of increasingly sophisticated experimental tools to detect and analyse microscopic events, new methods of mathematical analysis and computational advances have further accelerated the development in this area.

The discovery of dynamical chaos has opened up many exciting possibilities with wide-ranging implications in natural sciences and engineering sciences, CES being no exception. The usefulness of chaos comes from the fact that it is a collection of many sets of orderly behaviour, none of which dominates under normal circumstances. At first sight an engineer shies away from chaos, because he finds it unreliable, uncontrollable and, therefore, unwanted. However, the challenge for an engineer is to control chaos and make it manageable, exploitable and, in fact, invaluable.

We ought to recognize that the special problems encountered in CES have to be viewed within a grand universal framework. Indeed, forms of quantitative universality have been found to occur in chaos so that it

now seems that studies on a case-by-case basis may not really be necessary. Thus, for example, the routes to chaos by the system when a control parameter is varied follow specific mechanisms like period-doubling, intermittency or crisis. In each of these cases the onset has a typical diagnostic measure. For instance, in the case of period-doubling bifurcation, the ratio of successive bifurcation points tends to converge to a universal constant, now popularly referred to as the Feigenbaum number. The realization that the behaviour of complicated physical systems can be analysed on the basis of very simple nonlinear equations indeed suggests that the complexities of the real world may arise from underlying simple causes. Furthermore, for a system exhibiting chaotic dynamics, an important property is its sensitivity to its initial conditions. Even an infinitesimally small difference in specifying initial conditions accurately may lead to a catastrophic growth in the error. This means that one can make near-accurate short-term predictions about the system behaviour but certainly not long-term. This has profound significance in many CE systems that we deal with.

Multidimensionality, intermittency and coupling are conditions that are frequently met in many CE systems. Therefore, such systems become serious candidates for exhibiting chaotic behaviour. The subtleties brought out by the chaos theory pervade a number of situations, covering complex reactions in a CSTR, bubbling from a submerged orifice, mixing and so on. Let me illustrate the point further by giving some illustrations.

Chemical systems provide excellent prototype systems for studying nonlinear behaviour, since they offer a variety of convenient settings ranging from the simple to the more complex[8]. In this context, an open, well-stirred system (CSTR) following nonlinear kinetics and involving a few chemical species forms a good basis for the study of low-dimensional nonlinear dynamics and steady-state behaviour. Common examples of systems with nonlinear kinetics involve reaction mechanisms with autocatalytic or nonisothermal feedback features[9]. It is interesting that an early experimental observation of chaotic dynamics was reported by Schmitz *et al.*[10] for the now famous Belousov–Zhabotinskii oxidation reaction scheme. Continued theoretical and experimental investigations have revealed newer and newer nonlinear dynamical and bifurcation patterns. Attention continues to be given to analysing the behavioural patterns of systems of interest to chemical engineers, say, for example a nonisothermal consecutive reaction $A \rightarrow B \rightarrow C$ occurring in a CSTR[11,12].

Open systems operating far from equilibrium produce a very rich behaviour. Advances in modern instrumental methods such as photoelectron, field electron and field ion microscopy reveal the ensuing structures with higher and higher resolution power. Oscillatory, chaotic and spatiotemporal concentration patterns of the adsorbed species for carbon monoxide oxidation on platinum surfaces for various types of catalysts, including supported wires and foils, are available[13]. Specific domains of the catalysts show synchronization as a result of heat conduction through the supporting material or even through the gas phase. Propagating reaction–diffusion fronts (diffusion coupling) also leads to such synchronization. The important keywords are, therefore, coupling and synchronization.

Due to the coupling between complex kinetics and transport processes in CE systems, a wide range of dynamic behaviours can be observed. Possibility of dynamic pattern formation by coupling at a larger scale was demonstrated by Arce and Ramakrishna[14]. The dynamic behaviour and the stability of such systems and its linkage to the understanding of formation and propagation of long-range dynamic patterns (including travelling waves in reactors) is obviously a hot topic for study. True CE culture is entering into the field; the contributions by Fox and Villermaux[15] on the role of micromixing (a typical CE paradigm) in controlling the existence and nature of oscillatory patterns is a good example.

The development and study of nonlinear systems has been confined to study of low-dimensional systems so far. But a CE scientist has to look far beyond this. He has to deal with systems involving time delays (something that an electrical engineer may not experience in his systems) or those that are modelled by a set of partial differential equations. The time delay elements, for instance, are inherent in many CE operations, such as in feedback loops of well-mixed processes. The effect of their presence in the observed dynamics of a nonlinear system can be quite dramatic and they require attention.

Consider the classical problem of bubbling. We have been looking at this problem for decades. But suddenly a potentially new way of looking at it has emerged. Tritton and Egdell[16] found that bubbling from a submerged orifice is sometimes regular and sometimes irregular. They observed a period-doubling sequence leading to chaotic behaviour. Subsequently, intermittency of varying degrees and period-3 patterns were observed. Examples such as this suggest that such new tools can lead us to look at dispersion science in a very different way.

Consider another classical problem of mixing. Recent experimental studies and the merging of kinematics with dynamical systems and chaos are providing a paradigm for analysing mixing from a universal viewpoint. This is as well, since there is a large spectrum of problems studied in various disciplines in which mixing assumes importance (Figure 1). A universal framework can be used for analysis in different disciplines, whether it is mixing in a screw extruder or mixing in the mantle of the earth[17].

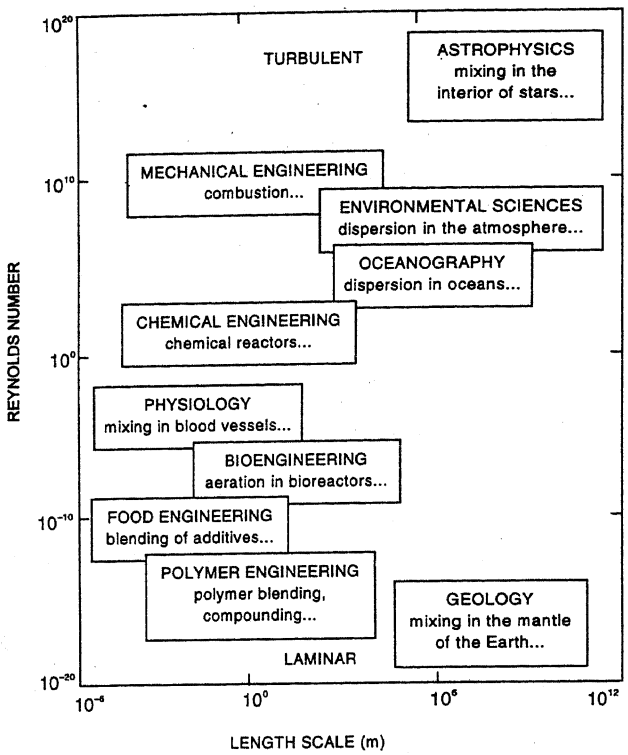


Figure 1. Mixing phenomena studied in diverse disciplines.

The real excitement lies in the potential of achieving laminar mixing of viscous fluids in a regime of deterministic chaos produced by intermittent change of initial conditions[18–21]. Here mixing results from a fundamental kinematic property. In such chaotic processes, mixing efficiency is not directly related to energy dissipation. The prospect of devising entirely new mixing devices based on these new principles is truly exciting. The recent work by Howes *et al.*[22] is interesting from this angle.

The examples I cited are typically those that are at the heart of chemical engineering. But let me move on and tell you about a more complex system, which combines the synergistic efforts from diverse disciplines being interpreted using the theory of chaos, which leads to new ways of taste detection. Let us consider membranes with biological activity. The central question concerns the interrelationship between the microscopic and macroscopic membrane properties. Properties such as the action potential, spontaneous firing and chaotic behaviour can be viewed now from a system-dynamical viewpoint. Thus, for example, the mechanism of oscillatory electrical potential of dioleyl phosphate millipore filter membrane placed between two solutions of different electrolyte concentration can be related to the macroscopic behaviour of interfaces generated due to chemically driven surface tension effects[23]. The system, when subjected to sinusoidal alternating current, synchronizes with the forcing frequency and also desynchronizes for changing initial conditions. Aperiodic behaviour in the form of chaos appears in the regime[24]. The appearance of chaos follows the route of quasiperiodicity and analysis of dynamic chaos enables us to use it as a sensor for detecting taste. The simple philosophy followed here is that the substances as received by biological membranes change the membrane potential, leading to sustained oscillation and spatiotemporal firing patterns of the neurons, which when transmitted to the cortex and the brain recognize the taste. It is thus possible to use an artificial lipid membrane to convert chemical information to electric currents and construct sensors that identify the taste. The prowess of the synergy of diverse disciplines is at once obvious here. A CE scientist should simply relish the idea of exploring, analysing and exploiting analogous phenomena and enriching the profession.

Fractals is another concept that has excited the physics community over the past two decades. Fractals are scale-invariant geometric objects with the property that when a small piece is magnified it resembles the whole object statistically. Many physical processes produce phenomena that show self-similarity over a wide range of scales. The methods of fractal geometry have been extensively used in diverse disciplines, ranging from atmospheric sciences to modelling of earthquakes and also in miniature systems in cell biology. Aspects of fractals in chemical engineering problems that we deal with, be they porous electrodes, catalytic surfaces, membranes and so on, have been dealt with very elegantly by Villermaux[5]. A lack of quantitative connection between the measured value of the fractal dimension and the underlying physics of the process has many times led to a scepticism about its meaningfulness. If one comes out with a fractal dimension after a lot of sweat and hard work, we often have someone saying, 'so what?'. Fractals are not merely physical objects but are strange attractors of the nonlinear systems, which make them all the more beautiful, while being all-pervading at the same time. The connection of physical and mathematical fractals is, however, important, since it links the nonlinear system dynamics to fractal geometry.

The methods of fractal geometry when applied to data obtained from mathematical model, while providing an insight into the behaviour of these nonlinear systems, are grossly inadequate for the measured experimental time series data. Also the presence of noise in the experimental data seriously distorts the results. Clearly, there exists a need for efficient noise reduction methods for dynamically sensitive systems, which need to be tested possibly on model examples first and then on real experimental systems. As more CE scientists get attracted to fractals, these words of caution may be useful.

CE scientist as a molecular information engineer

Molecular chemistry that a chemical engineer conventionally learns is that of a covalent bond. However, it is supramolecular chemistry that is holding the centrestage now. Supramolecular chemistry is the *designed* chemistry of the intermolecular bond[25,26]. It deals with chemical species that are held together and organized by means of intermolecular (noncovalent) binding interactions. The chemical, physical and biological features of these chemical species constitute the essence of this highly interdisciplinary field. It involves the organic synthetic techniques of receptor construction, metal-ion–ligand complexes involving coordination chemistry, studies of interactions involving physical chemistry and finally biological processes involving substrate binding and recognition. Its location at the intersection of chemistry, biology and physics makes it a unique one and it also offers great challenges to a CE scientist.

Designing systems for storing information in the structural (and eventually temporal) features of molecules and supermolecules constitute the real essence of supramolecular science and engineering. Supramolecular science deals with the process of molecular recognition. The essence of molecular recognition is that the molecular information is stored in the interacting partners. Chemical ingenuity is used to design receptors possessing steric and electronic features. These are designed to be complementary to those of the binding substrate. Further molecular manipulation provides a balance between the rigidity and the flexibility suitable for the function to be performed. Such subtle organization of molecular, structural and dynamical features forms the basis of *molecular information engineering*.

Why should a CE scientist get involved with this field? This is because molecular information engineering will be finally concerned with the design of molecular devices and of systems displaying higher forms of molecular and supramolecular behaviour such as self-organization, regulation, cooperativity, communication and replication. Let me explain this a little further.

We know that an interplay between molecular self-organization and molecular recognition of individual constituents leads to the construction of natural supramolecular systems. Such systems combine order and mobility. Their functioning depends on the organization. If CE scientists abandon the limits of their classical thinking, they can start conceiving the development of new materials, which simulate natural supramolecular systems, where their functioning will be governed by their organization. The connecting links between apparently separated disciplines of materials science and life science bringing in the central theme of supramolecular systems is demonstrated in Figure 2.

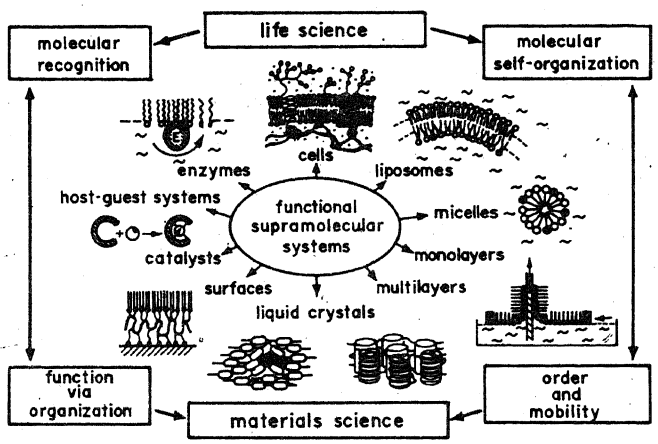


Figure 2. The connecting link between life science and materials science – the role of functional supramolecular systems.

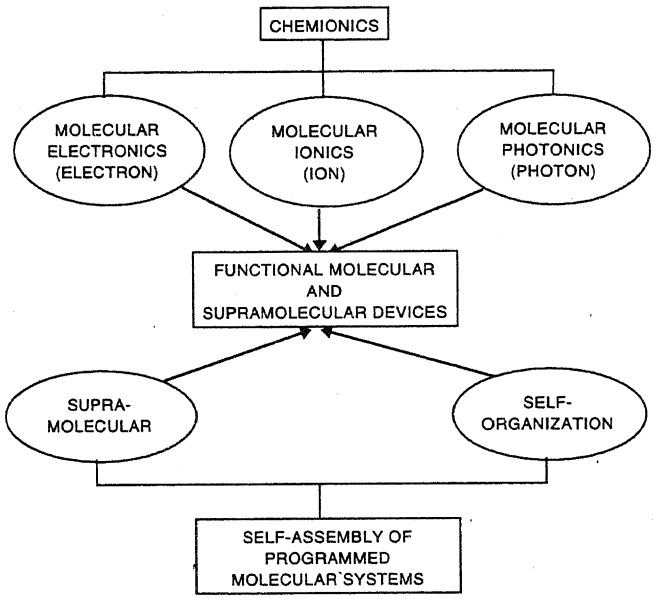


Figure 3. Chemionics as the chemistry of recognition-directed and self-organized photonic, electronic and ionic molecular and supramolecular devices.

But how do we enter this exciting field? The chemistry of molecular signal generation, processing, transfer, conversion and detection is already being hotly pursued by innovative chemists. The job of a molecular information engineer, if I may use that term, will be simply to design, build and operate molecular devices, which will be structurally organized and functionally integrated chemical systems built into supramolecular architectures. The function performed by such molecular devices will be the result of the integration of the elementary acts performed by the components, which may be photoactive, electroactive or ionoactive, depending upon whether they operate with photons, electrons or ions (See Figure 3). Strategies towards properties and

functions rather than towards structures will be the goal of a molecular information engineer. He will continuously seek to express a desired property in a molecule or a material, whatever be the composition and the structure.

What I have said so far sets out the general background. But now let me give a specific example of how CE scientists can use molecular information engineering in polymers to solve some challenging unsolved problems. I focus on the exciting idea of creating molecularly imprinted polymers. Molecular imprinting in polymers involves the idea of using a specific imprinting molecule to coordinate the assembly of synthetic monomers around a molecule of interest. The functional monomers are polymerized in the presence of a print molecule. The functionality in the print molecule interacting with the complementary functionality is found in the monomer units. A cocktail of functionalized monomers is 'prearranged' around the imprint molecule by noncovalent interactions, which could be ionic, hydrophobic, hydrogen bonding, etc. After polymerization, the imprint molecule is removed from the polymer. This leaves a polymer with recognition sites that are complementary to the imprint species with regard to both shape and functionality. This site will constitute an 'induced molecular memory' which will be capable of selectively recognizing the imprint species. The polymer will have a macroporous structure which will allow an imprint molecule diffusion into and out of the polymer matrix.

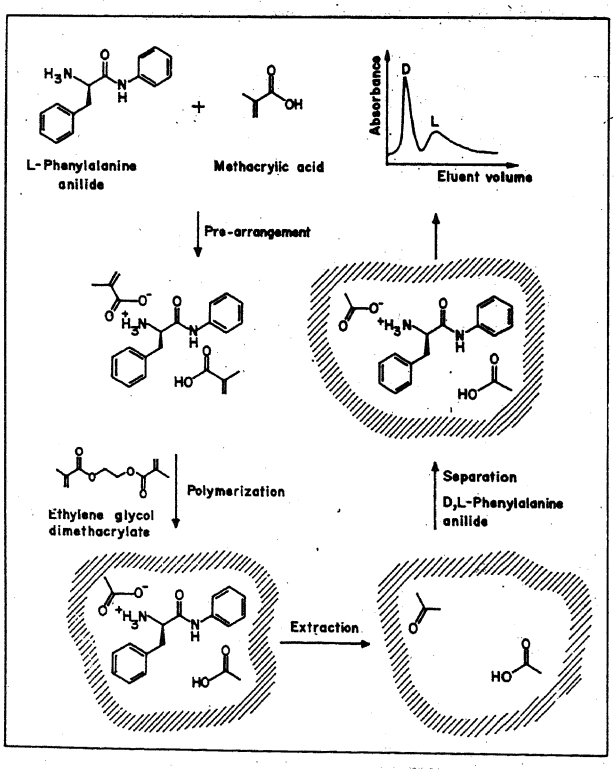


Figure 4. Schematic representation of molecular imprinting by a noncovalent approach of Mosbach[27]. An amino acid derivative is used. Different stages in imprinting technique leading to a molecularly imprinted polymer (MIP) capable of baseline separation of racemic mixtures are shown.

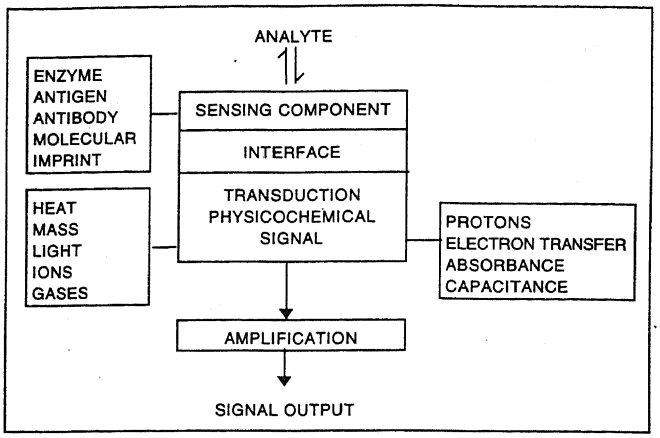


Figure 5. The principles underlying conventional biosensor operations and the possible role of MIPs in their function.

Molecularly imprinted polymers (MIPs) are exciting, since they have a fascinating range of applications. Figure 4 shows a schematic representation of molecular imprinting through a noncovalent approach using an amino acid derivative demonstrated by Mosbach[27]. Here, after mixing the imprint species with functional monomer and crosslinker, polymerization is initiated and the print molecule extracted from the resultant polymer under acidic conditions. Such MIPs find immediate applications in the challenging task of separation of racemic mixtures. Separation factors of between 2 and 8 have been achieved. What is impressive is that the molecular memory of the MIPs is not lost despite their manifold use in racemic resolutions over a period of several months.

This opens up an exciting new field of separation science for us in a challenging area such as chiral separations, where so far traditional chiral stationary phases (CSPs) have been used. Tailor-making binding polymers for a given separation process without involving the 'trial and error approach' inherent with different CSPs will be a big gain if we followed this strategy. The science–business connection is obvious here too. Please remember that of 500 or more optically active drugs presently in the market, about 90% are administered as racemic mixtures. Food and Drugs Administration regulations now require that for any new optically active drug, both enantiomers must be treated as separate substances in pharmacokinetic and toxicological profiling. Molecular imprinting might be the way to permit enantiopolishing, i.e. the removal of small amounts of contaminating optical isomers on a commercial scale. Separation and purification of wide ranges of molecules by exploiting the stereospecificity and regiospecificity

promises enormous potential with applications in differentiation of isoenzymes, conformers or epitopes and so on.

One can extend this molecular information engineering further, since MIPs can also move into substrate-selective polymers as sensory components. The idea will be to substitute the bio part in an enzyme- or antibody-based sensor with catalytically active or ligand-specific polymers prepared from specific polymers imprints. These can then move towards more robust sensing element systems (Figure 5).

In many cases no suitable biomolecule will be available. It is here that molecular imprinting will have the greatest potential. It will be able to create the custom-tailored synthetic binding site for a given molecule. The possibility of 'direct contact' systems, where 'molecular wires' could be created linking the binding cavity directly to a transducer, such as an electrode, to give rise to an instant signal upon binding follow naturally. A whole range of exciting possibilities will open up as soon as the real prowess of molecular information engineering is seen by CE scientists.

CE scientists and living systems

The fascination of chemical engineers for living systems dates back to several years. Whether it is transport phenomena[28] or reaction engineering[29] in living systems, CE scientists have brought elegant tools of sophisticated analysis to this field. Modern biotechnology, be it industrial, medical, environmental or some other, has always put a great demand on truly borderless behaviour between diverse disciplines. For instance, medical biotechnology thrived only because of a close cooperation between biologists, biochemists, geneticists, immunologists, medical specialists, chemical engineers and so on. Challenges of implantable and extracorporeal devices that duplicate the function of mammalian organs or facilitate the prevention or treatment of diseases are being met only through borderless laboratories combining several disciplines. Metabolic engineering and tissue engineering, in which CE scientists have got involved, require strong interdisciplinary teams. Search of novel biocompatible materials for prosthetic replacement and repair of both hard and soft mammalian tissue is also showing promising results only due to such synergy of action. Unresolved problems in industrial biotechnology of development of highly accurate, sensitive, and selective sensing devices capable of measuring concentrations of bioactive substances and metabolic components in bioprocess streams and monitoring these parameters for computer control also require strong interdisciplinary teams.

What chemical engineers have already achieved is something to be proud of. Significant contributions have been made in industrial biotransformation processes[30], the subject of last year's Danckwerts memorial lecture. Chemical engineers had the lion's share in the development of artificial kidney, pursued initially as a natural extension of their work on membranes. They were at the centrestage when insulin was produced on a large scale using recombinant DNA technique. Now they have moved a step further by developing hybrid pancreas. The idea is to take live pancreas cells and immobilize them in an inert matrix to be used as hybrid pancreas. In drug delivery systems, we are now talking of low flow pumps and smart membranes. In low flow pumps, microcapsules are so produced that each releases its contents at a specified time. Smart membranes will sense the glucose level of the blood and release the required amount of insulin. Attempts are even being made to generate artificial red blood cells, which will not contain any living matter and yet perform the functions of the cells in great detail.

In this context it is worth quoting from an editorial that was written by Frauenfelder and Berg[31] in connection with a special issue of *Physics Today* that dealt exclusively with the emerging new relations between physics and biology. They said 'interface between physics and biology ... is a cross-field endeavour to which physicists can make major contributions. But to do so, they need to learn enough biology to be able to talk with biologists and to distinguish problems that are central from those that are peripheral. Since much of biology is descriptive, the physicist has a strong advantage: It is much easier for a physicist to learn biology than for a biologist to learn physics'.

Just replace the word 'physicist' with 'CE scientist' and read the above paragraph again. The idea is clear. CE scientists have already played a crucial role in bringing quantitation to biology[32]. The tools and skills that we will use in future will also come from diverse disciplines again. My chemical engineering colleague in NCL, Bhaskar Kulkarni, recently wrote a paper[33] on the application of artificial neural networks for procaryotic transcription terminator predictor and it featured on the front cover of the journal *FEBS Letters*. The interesting point here is that a chemical engineer has picked up a tool developed by electrical engineers and mathematicians and contributed to quantitation of a biological problem!

CE scientists are making headway in developing more advanced methods. The growth of procaryotic cells has been predicted in a rigorous quantitative way – a far cry from our gross Monod and Michaelis–Menten equations. Such new models consider the formation of proteins from amino acids, RNA as well as DNA from nucleotides and cell wall membranes from lipids, not only for the parent cell but also for the plasmids contained in it. The distribution of plasmids as the cells multiply has also been predicted.

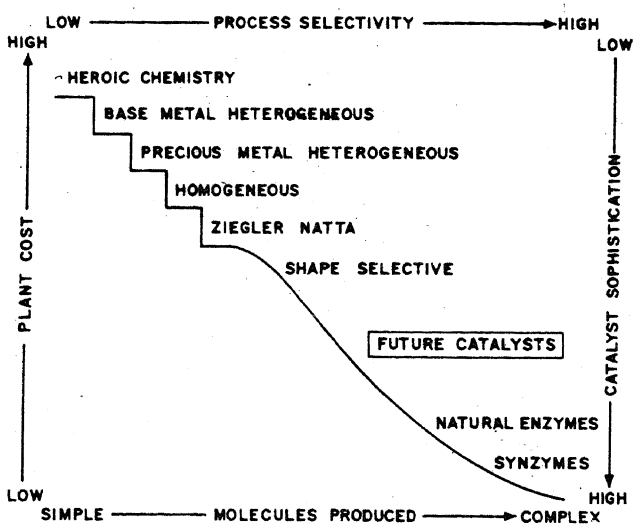


Figure 6. Historical trends in the development of catalysts.

Let us, for a moment, look at enzymes, which the CE community is looking at as catalysts of the future. Figure 6 shows the way in which catalyst development has moved over the years. The chemical engineers have so far contributed to the development of production processes of commercial enzymes. They have tried to immobilize them and so on. However, once again if they become borderless then they can get involved in an even more exciting game, namely displacing the enzymes themselves. Indeed, designing catalysts functioning in a manner similar to natural enzymes, i.e. creating synzymes, is emerging as an exciting challenge, as we can see from Figure 6.

Enzyme catalysis involves binding of the reacting substrate in a perfectly fitting cavity or cleft of the enzyme that contains functional groups in the correct stereochemistry for binding, catalysis, and group transfer. The idea of using zeolites as hosts has opened up the field of zeozymes. Use of molecularly imprinted polymers now offers interesting possibility for creating enzyme-like catalysts too. But as we become more and more borderless, we can start bringing in diverse ideas from different disciplines. A recent breakthrough[34] is worth describing.

Many attempts have so far been made to create a synzyme for the purpose of catalytic oxidation on lines of the enzyme cytochrome P-450. Heterogeneous mimics have been designed that use organometallic complexes encapsulated in the supercages of zeolites, which are the so-called zeozymes. However, these systems have no real mechanistic analogy with the enzymatic process, and the oxidation rates tend to be low. Parton *et al.*[34] tackled this problem by developing a composite

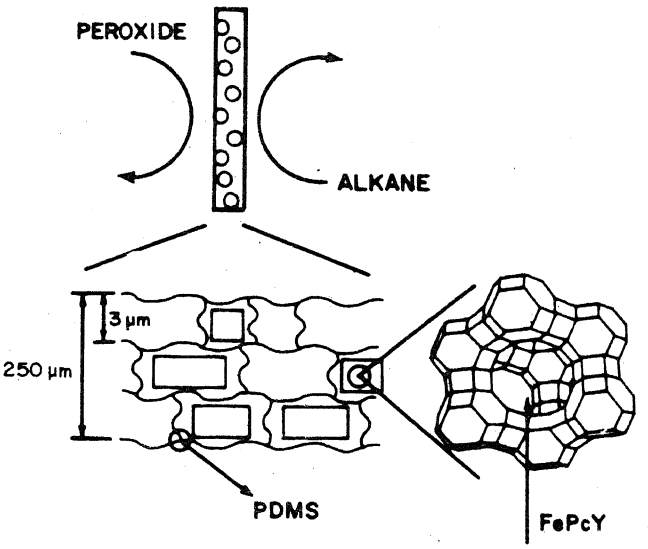


Figure 7. Zeozyme in a polymer membrane providing a mimic of cytochrome P-450.

catalytic system by incorporating iron phthalocyanine complexes encapsulated in crystals of zeolite Y, which were in turn embedded in a polydimethylsiloxane membrane (Figure 7). The polymer served three purposes. It acted as a mimic of the phospholipid membrane in which cytochrome P-450 resides. It acted as an interface between the two immiscible phases. It avoided the need for solvents or phase transfer agents. The system oxidized alkanes at room temperature and showed amazing turnover. The premium here again was on creativity, daring and combining concepts from diverse disciplines in an uninhibited fashion!

The way the field will move will be hard to predict by any logical and structural analysis. We already have examples of animals as bioreactors – take the case of transgenic cows producing human lactoferrin or transgenic sheep producing α-1 antitrypsin. Modern biotechnology is thus helping us to replace bioreactors by animals. On the other hand, modern polymer science is helping us replace animals by plastics! Let me explain this.

Immunoassay techniques require antibodies produced from laboratory animals. How about making plastic antibodies? This is not a dream anymore. All that we need is an exploitation of molecular information engineering done on chosen polymers so that they will become antibody mimics which will be comparable to biological

antibodies. It has been shown[35] that molecularly imprinted polymers could be made which show strong binding and cross-relativity profiles very similar to antibodies. Now this opens up an enormous opportunity, since MIPs can be made against a great number of organic molecules, for example, drugs, metabolites, hormones and toxins, provided they have suitable interaction points. It is easy to see then the amazing synergistic effect of diverse disciplines merging together.

CES and nanotechnology

Great expectations on nanosystems (1–100 nm) have been raised recently. Exciting pictures of nanoscopic machines that replicate, build crystalline rocket engines, repair cells and extend life spans have been painted[36,37] and even new strategies of doing all this by micromachining have been suggested. However, there is a problem here, since such nanomachines must themselves be put together by other nanomachines! Ideas of building molecules by pushing together atoms by the tip of the scanning tunnelling microscope have also been suggested. We can only imagine how slow and cumbersome these processes will be. Solid-state physicists, materials scientists, and electrical engineers build nanostructures through techniques such as chemical vapour deposition, etc. However, fabrication of microstructures and devices requires increasingly substantial effort as they are extended to the range below 100 nm. Then how can we proceed?

While we are all struggling, biology abounds in nanostructures with dimensions of about 1–100 nm (ref. 38). Just consider the range of biological structures – from proteins to viruses to cellular organelles. The challenge is to make structures of the size and complexity of biological structures, but without using biological catalysts or the information coded in genes. Nanotechnology, therefore, requires new strategies and technologies.

Self-assembly is emerging as a new strategy in chemical synthesis, with the potential of generating nonbiological structures with dimensions of 1–100 nm. Structures in the upper part of this range of sizes are presently inaccessible through the chemical synthesis route and the ability to prepare them would open a route to structures comparable in size (and perhaps complementary in function) to those that can be prepared by microlithography and other techniques of microfabrication.

What can CE scientists do here? One option is to look at the nature and see how we can exploit the self-organized structures that it has already created. The second is to understand the process of self-organization and simulate it synthetically to create new structures. Both approaches require ingenuity but both are possible, as I will show.

Surface layers (S-layers) represent the outermost cell wall component of many eubacteria and archaebacteria. These S-layers are arrays of protein or glycoprotein subunits that are linked by noncovalent interactions. Channels of 2–6 nm diameter exist in S-layer lattices. This beautiful S-layer material was isolated from highly thermophilic eubacteria, crosslinked with glutaraldehyde and then it was deposited on commercially available microfiltration membranes[39]. The only function of the microfiltration membrane was to provide mechanical support. These membranes displayed unusually steep rejection curves. Such new types of UF membranes will have considerable potential in opening up new doors to ultrafiltration technology.

Let us now turn to synthetic methodologies. Block copolymers have particularly interesting features with respect to equilibrium statistical mechanics, forming various phase-separated domain structures with nanometer periodicities, the so-called nanopatterns. Recently, a breakthrough was achieved[40] in using the process of self-organization in block copolymers of polystyrene–polyparaphenylene. The films, of 10–30 μm thickness, were produced by evaporating solutions in carbon disulphide under the flow of a moist gas. Empty spherical cells about 0.2–10 μm in diameter appeared spontaneously in a beautiful hexagonal array. Such membranes might find application in controlled release of drugs or other bioactive species, or as materials with useful optical properties, moulds or scaffolding for forming ordered microstructures, and model substrates for surface science. What is remarkable is that the polymer architecture is able to induce a complex regular morphology at a scale two or three orders larger than the molecular scale. The mechanism of formation of this new regular morphology involves a combination of complicated thermodynamics and transport phenomena and provides an interesting challenge to CE scientists.

Nanobiotechnology is emerging as an exciting field. Fabricating organized biomolecular structures is a new challenge that is being met by bringing in diverse disciplines. Genetic engineering will tailor protein's makeup and also its function. Organic chemistry will provide new materials for anchoring and preserving the altered proteins. Electrical engineering will offer ways to detect signals from internal working of proteins. We can finally turn a small assembly of biomolecules into a custom-designed molecular machine.

Another exciting area is where proteins are being turned into nanoscale chemical processing plants. Proteins can be used as reaction vessels for controlling the particle size of other materials. They can cage several compounds. If the chemical affinity of the protein can be changed, then the protein can be engineered to fit the mineralization product that one wants inside it. Such

caged particles will pass through the human body in an unrestricted way. They will be useful in diagnosing and treating diseases. It is clear that CE scientists can share the grand challenges of nanotechnology by joining hands with others working in protein engineering, organometallic chemistry, semiconductor technology, etc.

Entering new territories

Venturing into new fields will be both exciting and rewarding for CE scientists. Nontraditional thinking, whether it is in terms of creative use of ultrasound, photoenergy, microwaves or some other means, will continuously give the CE scientists new challenges and new opportunities. Laboratory curiosities or chance observations in other disciplines can be turned into major opportunities with winning technologies if we are both vigilant in watching other fields and daring. Let me take one specific example in detail, namely sonochemical reaction engineering.

Chemists have shown over the years that application of ultrasound yields entirely different products from the same reactants, enhances the rates of the reactions phenomenally and even permits at very mild reaction conditions synthesis of products that can normally be done only under very high temperatures. Although catalogues of the use of ultrasound by chemists are available[41], little or no work aimed at quantitative prediction of the effects of ultrasound on the course of chemical reactions has been carried out. Now sonochemical reaction engineering is beginning to emerge as an exciting area of research for CE scientists, since it combines so many novel elements of cavitation dynamics, transport phenomena and reaction engineering.

The influence of ultrasound on chemical reactions is known to come from the cavitation phenomena. Very small bubbles of dissolved gases that are omnipresent in the liquid can grow in the rarefaction cycle of the sound wave and collapse rapidly and adiabatically during the compression cycle. The adiabatic compression generates very high temperatures and pressures in the bubble. Reactive intermediates are then born, which can diffuse out and react with other reactants present in the surrounding liquid phase. Further, when a cavitating bubble collapses near a solid surface, it does so asymmetrically and generates microjets of high velocity. The resulting microjets can clean the solid surfaces. Such a cleaning action leads to reactivation of passivated catalytic surfaces and also local heating of catalytic surfaces. This results in impressive enhancement of the rates of reaction. The high velocities generated can also decrease any diffusional resistance present and, once again, enhance the rates of the reactions.

The big question is how can sonochemistry be taken through the path of sonochemical reaction engineering to sonochemical technology? A body of engineering knowledge has still to be acquired before design of reactors or even scale-up can be achieved. The first paper on sonochemical reaction engineering, which provides a high level of predictive ability, appeared a few months ago[42]. It is interesting to see that a multidisciplinary team of chemical engineers, mechanical engineers and chemists worked together to provide insights into the fluid-dynamical phenomena as well as the reaction events.

A lot more needs to be done in sonochemical reaction engineering. Capturing the dynamics of a cavitating bubble, postulating the rates of reactions occurring in the bubble that produce reactive intermediates, modelling of convection created by microjets and acoustic streaming, investigation of pitting and its effects on creation of new surfaces, understanding the physics of the fusion of particles and the new catalytic chemistry are some of the interesting challenges in modelling. But an even bigger challenge is to make a sonochemical reactor work in practice. New ideas on loop reactors are emerging[43]. Normal sonicators have the problem of erosion of the power-transmitting surfaces due to cavitation. How do we concentrate the power away from the transducer surfaces? The use of spherical vessels for batch reactors or hollow cylindrical-type transducers to position the power around the centre of the axis seems to be a possibility. However, only a close interaction with acoustic engineers, mechanical engineers, chemists, etc., can move this field further.

The use of ultrasound is just one possibility. Interesting opportunities exist for the development of applications of both electrochemistry and photochemistry. Unusual cross-disciplinary thinking is needed though. The concept of using electrical potentials impressed upon metallic or semiconductor catalysts to bias either the rate or pathway of a heterogeneous catalytic reaction to affect the reaction kinetics or product distribution is an unexplored area. Can we carry out electrochemical (i.e. electron insertion/withdrawal) processes at electrode surfaces in the proximity of immobilized enzyme systems as a means of conducting oxidative or reductive enzymatic transformations? This will do away with the necessity of cofactor reactants. Another interesting challenge is the simultaneous generation of electric power with production of useful organic intermediates by controlled catalytic oxidation or reduction at an electrochemical cell electrode. Such 'bootstrap' chemical processes, which are energy- and raw-materials-conserving, will make the profession leap-frog.

If we go really borderless, then we can think in terms of the combined-fields approach[44]. We can generate additional driving forces in a process that works simultaneously with the conventional driving force of the process. The additional forces may evolve from the application of external fields or body forces. Electric

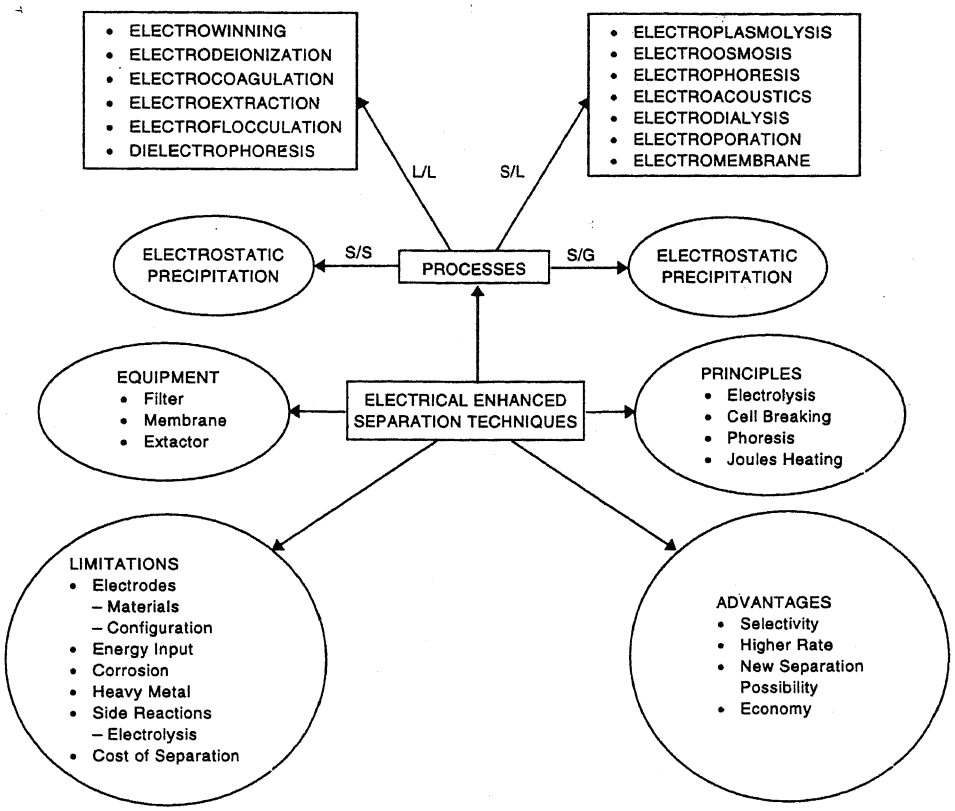


Figure 8. Range of electrical enhanced-separation techniques.

field, which is one possibility, has already been explored for enhanced-separation techniques (Figure 8), but my feeling is that we have only been scraping the surface so far and a lot of unconventional thinking has to go on here.

Take, for instance, an imaginative use of the electrical field in the pulsing mode. We know[45] that an electric pulse of intensity in kilovolts per centimeter and of duration in microseconds to milliseconds can cause a temporary loss of the semipermeability of cell membranes, thus leading to ion leakage, escape of metabolites, and increased uptake of drugs, molecular probes and DNA by cells. This phenomenon is referred to as 'electroporation'. Electroporation can find applications in insertion of proteins into cell membranes, improving drug delivery and hence effectiveness in chemotherapy of cancerous cells, activation of membrane transporters and enzymes, etc. Why not design very cheap disinfection or water purification processes based on electroporation?

Let us pursue the idea of pulsing and look at other possibilities. Consider membrane fouling during cross-flow microfiltration. We can imaginatively use electrolytic membrane cleaning. Here again, intermittent electric field pulses can be used to clean membranes without the interruption of the separation process. It has already been shown[46] that this technique can improve the overall membrane permeation rates by a factor of up to 10 with only a modest electrical energy requirement. Another technique, namely the *in situ* electrolytic membrane restoration, is very effective in restoring membrane water permeation rates under mild chemical conditions. Used in combination, these techniques have the potential of obviating the requirement of membrane backflushing and conventional chemical cleaning.

Let me emphasize again that these examples are by no means exhaustive but are only typical of the opportunities that exist. Photochemical and electrochemical reaction engineering have already matured. Sonochemical and microwave reaction engineering are emerging as new disciplines. Asimov's orbital reaction engineering[47] may be a long way away – but emergence of other fields will depend on our creativity and daring again.

New look at old materials through seamless science

With rapid advances in science, a CE scientist will have an opportunity to look at old materials and make them perform new functions. He may even convert them into smart materials with attributes of selectivity, sensitivity, shapability, self-recovery, self-repair, self-diagnosis, self-tuning and switchability. Use of some severe interdisciplinary science is required here to make this happen. For instance, 'dumb' water-swollen gels, which are

just used as passive supports, can be so transformed that they bend, walk and even mimic enzymes! To show such an exciting range of possibilities, I will just choose three familiar materials, namely gels, zeolites and clays.

Let me focus on gels first. Biological materials in the body are usually composed of soft and wet materials. This is in contrast with most industrial materials that we chemical engineers use. Take, for instance, metals, ceramics and plastics. These are dry and hard. Seldom do we use soft materials. But it is beginning to emerge now that 'soft and wet' materials can be important not only as biological materials that construct the body but also for displaying their functions. Designing mobile machines using soft and wet materials, furnishing soft materials so that they are able to make shape changes, or to generate tensile stresses that can lead to motility without the requirement of a rigid structure are challenges that will be met in the new world of seamless science.

Let me discuss more specifically polymeric gels, which are swollen polymeric networks. They possess the cohesive properties of solids and the diffusive transport properties of liquids. The combination of elastic properties and osmotic reactivity makes them very special. Gels are held in place by a hydrophilic lattice of long-chain molecules. The fragile balance between the electrostatic repulsion of the adjacent polymer strands and the osmotic pressure exerted by the ions in the solvent holds the key to the behaviour of gels. Disrupting a gel's charge distribution by altering its environment can trigger it to swell or shrink suddenly[48,49]. Such phase transitions can be exploited imaginatively.

The thermodynamics of phase transitions, aspects of transport phenomena as well as the energy conversion phenomena in these gels will be best understood by CE scientists. Indeed, the entry of CE scientists in this area is giving rise to the development of several systems of interest, including separation technology[50,51], sustained release systems[52] and so on.

It is rather interesting to notice the simultaneous generation of ideas in such fields that take place between different disciplines driving different applications using similar principles. Our research group recently used the pH sensitivity of the hydrogels for attaining time-invariant diffusional flux through bilayered membranes with collapsible barriers[52]. The novelty of the approach was that the bilayers were formed *in situ* and the diffusional behaviour was manipulated at will depending on the external stimuli. The idea rested, however, on making both surfaces of a gel slab collapsible. Our Japanese friends[53] decided that they would collapse only one of the surfaces of the gel. Their idea was to create a bending motion (Figure 9). By exploiting this asymmetric collapse with the help of a rachet-type 'gel looper', they showed that a gel could 'walk' along a bar! As chemical engineers, it is easy to see that we were interested in controlling diffusion and the Japanese were interested in soft robots – but the phenomena exploited by us were exactly the same.

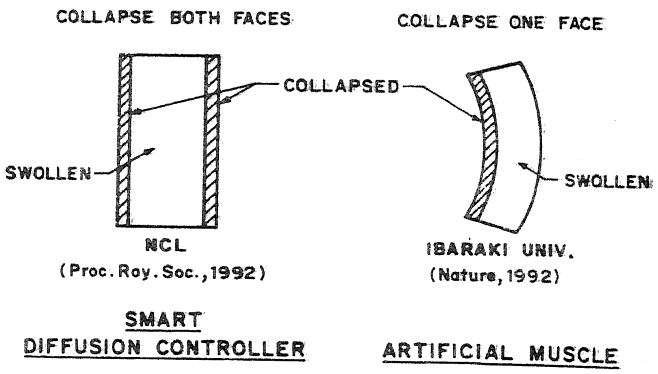


Figure 9. Schematic presentation of two applications of a reversible collapse of a gel bar surface used for designing a smart diffusion controller[52] and a mimic of artificial muscle[53].

In contrast to motors and hydrodynamic pumps, the motion of a polymeric gel demonstrated by Osada *et al.*[53] is produced by the chemical free energy of the polymer network. Thereafter, electrical or thermal energy is used to drive the direction and control the state of the equilibrium. Thus, the chemomechanical gel driven by outside stimulus is able to exhibit 'gentle' and 'flexible' action, and its movement is more akin to that observed in a muscle than in the metallic machine systems. The application of chemomechanical gel as an 'actuator' for robots and other systems has special interest. The molecular machine that is capable of biomimetic motility by using gel will allow entry into a new world of movement where the system works smoothly and gently without noise, heat evolution and exhaustion[48, 54]. Applications might eventually include the construction of devices that can be inserted into muscle and even emulate the nervous system. This field opens up so many new exciting vistas for CE scientists.

Let us now move to zeolites, which were initially discovered as naturally occurring materials. Much zeolite research and development with which chemical engineering scientists are associated is still aimed at the traditional uses such as catalysts and adsorbents. However, over the past few years we have seen zeolites emerge as advanced materials[55]. Indeed, zeolites as microporous materials with nanometer dimension window, channel and cavity architecture present new opportunities. The key in assembling and maintaining controlled microstructure is to use solid host lattices that serve as templates within which a guest structure of nanometer architecture can be assembled. This opens up new possibilities.

Take the field of solar energy conversion. Here a spatially well-organized molecular system is required that models photosynthetic mechanisms. This will enable

redox reactions to proceed efficiently along the desired paths. Other potential applications include zeolite electrodes and electron relays, zeolite batteries, zeolite fast-ion conductors, intrazeolite semiconductors, zeolite chemical sensors, zeolite imaging and data storage materials, zeolite lasers and displays, as well as zeolite composites to form perm-selective membranes or thin films.

The stage is set for rapid developments of zeolite-based electronic, optical, magnetic and dielectric materials. Zeolite semiconductors may not be far away. The marked proton conductivity of certain acid zeolites may open up the ways to electrochromic zeolites. The emergence of acentric zeolites may lead to the design of zeolite piezoelectric devices. The molecule-sized pore and channel structure of zeolites could provide a novel medium in which to create ultrahigh resolution images and package and process information at high storage densities. The unidimensional channel architecture of certain zeolites may even permit the assembly of low-dimensional zeolite conductors and the control of energy transfer and energy migration in restricted dimensions. How do we overcome the present disadvantages of zeolites, namely pore plugging, poisoning, structural defects, the difficulties of growing large single crystals, and interfacing a zeolite device with the outside world? A CE scientist is well poised to take upon himself some of these challenges. But he must operate in the borderless way interacting with materials scientists, physicists, electronic engineers and so on.

Let us look at another old material that we are so familiar with – namely, clays. There have been suggestions that clays may have served as templates for the first self-replicating systems on this planet[56]. Indeed, clays have been considered as possible chemical microreactors in the prebiotic synthesis. They may very well have served as 'low-tech' precursors to the self-replicating biomolecules. Oligomers of glycine have been formed in the presence of clays subjected to cycles of wetting and drying simulating natural cycles. It has been demonstrated that energy due to mechanical stress could be stored within the lattice as trapped holes, and released in the form of additional acidic catalytic sites. Luminescence can be induced by dehydration of a clay and mechanical stress can translate into triboluminescence. In short, clays can store energy from the environment and release it in various forms.

Is it not then possible to let our imagination run wild and aggressively move into this area? We can raise a number of questions. Can clays oligomerize simple molecules into biomolecules? Can the required dehydration reactions occur in an aqueous environment? Can clays protect their reaction products from photodecomposition under intense irradiation from sunlight? Can clays store energy and release it by transforming it into chemical energy to operate these condensation reactions? Can clays distinguish stereoisomers to form certain products selectively? The answers to most of these questions are positive. Can the CE scientist not move aggressively to exploit these materials?

What I have tried to emphasize here is that innovation is the key to making mundane materials smart or adding enormous value to them. However, the innovation will emerge only if we intermesh intimately with other adjacent sciences.

Acquiring new tools from other disciplines

We must empower a CE scientist with sophisticated tools, both experimental and computational, for him to march ahead. The quest of a CE scientist to probe molecular-level events and structure on the one hand and to gain access to the spatially resolved macroscopic processes on the other will mean that he will be in constant search of new tools. Especially valuable are those tools which allow the measurement of the dynamics of the events (say during mixing) or allow *in situ* measurements during say a reaction–diffusion event on a working catalyst. Once again these tools will be available elsewhere and the real ingenuity will lie in locating them and using them. Thus, tomography technologies are extensively used in medical diagnostics. A thermal neutron-based Computerized Tomography (CT) system may be used conventionally for industrial diagnostics. It is up to the CE scientist to know the relevance and the worth of these tools and bring them into the profession.

If we look around, we see that NMR spectroscopy and imaging, tomography, trajectography, image processing capability, etc., are serving as powerful aids for CE scientists, who are searching for more and more detailed information. Let us see how this is being achieved by taking a few specific examples.

Tomography techniques used in the medical field cover a wide range of tomographic sensing techniques. These include nucleonic (X-ray, γ-rays), optical, ultrasonic and electrical methods. Some of these methods would have limitations for CE problems at hand. For instance, optical sensors are not useful in opaque systems. Radiation-based sensors are slow, expensive and have safety constraints. However, the scope for the use of these techniques is enormous. Electrical capacitance tomography can be used for nonconducting systems. Electrical resistance tomography can be used for conducting systems. Optical tomography has already been used for imaging mixing process within agitated vessels.

An eye for application of new tools will take us a long way. For instance, consider positron emission tomography (PET), an imaging technique widely developed for diagnostic medicine. It was imaginatively applied by chemical engineers to examine the problems of mixing of powders and gain some valuable new insights[57,58].

Similarly consider a commonly used technique in industrial diagnostics, namely thermal-neutron-based CT system. It can be operated in a radiograph mode to obtain 2-D transmission images or in the real-time mode to image dynamic events. Both systems can be used in conjunction with computerized tomography for 3-D imaging. Jasti and Fogler[59] used this technique imaginatively recently in getting some important insights into porous media flow.

Let us cite a specific example to illustrate the power of these techniques. Williams *et al.*[60] exploited electrical impedance tomography to study mixing in stirred vessels. The batch mixing of a slug of brine tracer was mapped out. Simultaneous capture of video images of the tracer mixing provided a visual confirmation of the tomography reconstruction. The experimental acquisition of fully 3-D concentration field information was particularly satisfying. Such work provides a powerful incentive to develop and implement fully 3-D models of the flow and mixing, since the details available now are precisely commensurate with the details that are capable of being provided by computational fluid dynamic (CFD) models.

CE scientists getting involved in such endeavours must not be just at the periphery but must also get involved in obtaining deeper insights into image reconstruction problems connected with spatial nonlinearities and also the utilization of 3-D data. Furthermore, after a while, it is also important that they move from the 'comfort zone' of studying 'clean model systems' to 'real-life systems' with these powerful tools. The idea is to solve the 'problems that need to be solved' rather than those that 'can be solved'.

Let us move on to NMR techniques now. Excellent reviews by Gladden[61,62] show the great power of this tool in chemical engineering. NMR probes physical and chemical phenomena occurring over a vast range of both length and time scales. While NMR spectroscopy yields information on structure and dynamics at the molecular level, NMR imaging probes macroscopic phenomena such as structural heterogeneities within materials, filtration processes, concentration and velocity profiles, etc. Such studies enable additional insights into the physics of the problems and may even suggest new strategies.

We ourselves have used NMR extensively over the past few years and found it to be a very versatile tool – for probing both microscopic as well as macroscopic events that take place during the swelling and dissolution of polymers[63–65] as well as during the hydration of superabsorbing polymers[66–68]. Let me just illustrate the point by giving an example. We looked at the molecular-level events while studying the dynamics of mobilization in swelling-dissolving polymers. We saw strange plateau appears in the averaged molecular mobility vs. time curves (Figure 10 *a*) as well as in fraction of

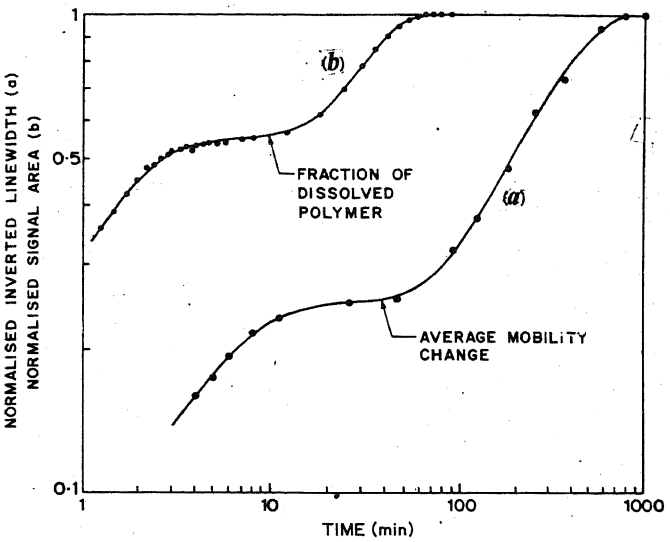


Figure 10. *a*, Reciprocal olefinic proton line width reflects on the average mobility in a swelling–dissolving polyisoprene polymer. The non-monotonous increase of mobility with time is striking[65]. *b*, Normalized signal area measured during dissolution of high molecular weight polystyrene reflects on the dissolved fraction of polystyrene in $CDCl_3$ solvent. The plateau is striking and is influenced by disengagement dynamics[64].

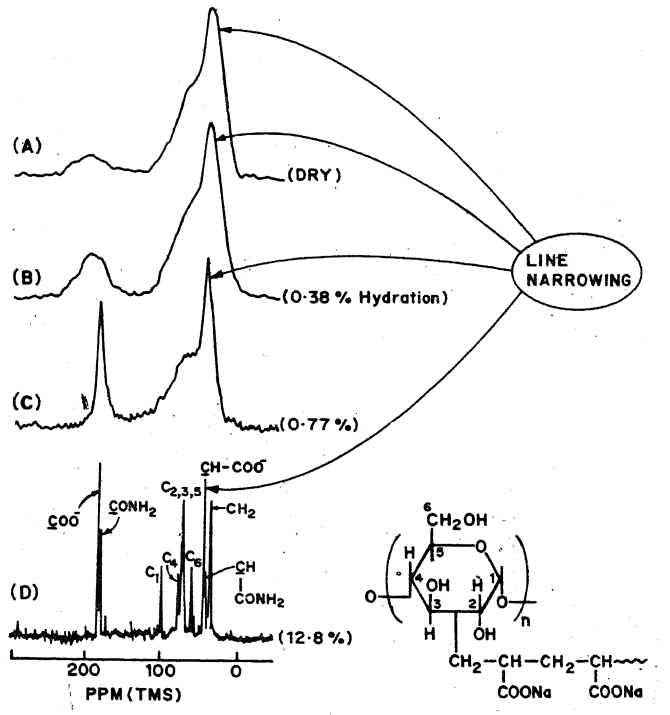


Figure 11. Specific hydration of a superabsorbing polymer (with the structure shown) studied by high-resolution solid-state NMR by Rajamohan *et al.*[68]. Line narrowing at the carboxyl shows preferential hydration at that site despite other hydrophilic groups present in the polymer.

polymer dissolved vs. time curves (Figure 10 *b*). These unexpected observations helped us understand the role of disengagement dynamics at the gel–fluid interface for the first time. We have even been able to examine specific hydration sites in superabsorbing polymers (Figure 11), showing how water mobilizes the carboxyls preferentially, first in preference to other hydrophilic groups. Thus, NMR is a powerful tool that gives important insights into both macroscopic and microscopic phenomena.

In the above example, although we studied molecular mobility of macromolecules, it was still averaged over the sample dimension. However, some recent spectacular advances enable us to observe the dynamic events of reptation of a macromolecule even at a single-molecule level.

Perkins *et al.*[69] observed the tube-like motion of a single flexible DNA molecule by using fluorescence microscopy. Kas *et al.*[70] reported direct observation of reptation obtained by video microscopy of fluorescent labelled filaments of actin in a solution. How such improved understanding of molecular processes of migration and reorientation of DNA molecules undergoing conventional and pulsed field gel electrophoresis, as obtained by direct observations from a microscope and computer modelling studies, has helped in improving these sophisticated separation processes has been shown by Bustamante *et al.*[71]. Advent of such studies at molecular level will empower CE scientists to examine problems of interest to them with even greater rigour.

Let us move from experimental tools to computing tools. CE scientists have not been found wanting in moving quickly and exploiting new advances, be they in artificial intelligence (AI) techniques or in parallel processing, or in some other areas. They have not caught on to virtual reality as yet, but I am sure they ought to – and they will indeed do it soon.

A recent AI application in chemical engineering is the use of artificial neural network (ANN). ANN is a numerical (not symbolic), empirical modelling tool that has received much attention recently[72]. The ANN developments began with computer modelling of the human brain with a view to emulate neuron interactions. The universal approximation characteristic of ANN makes them extremely useful for applications such as on-line fault diagnosis, process control, process modelling, process and product design, trend analysis, etc. The ground is now ripe for wide-ranging applications. The ability to adapt, learn from noisy, incomplete or inconsistent data, better filtering capacity and true multiple-input–multiple-output nature make them indispensable in modern times. In most situations in real-world applications, a chemical engineer would neither have a formal model nor complete experimental data and ANN can help him greatly. Besides serving as a numerical modelling tool in intelligent process manipulations, when combined with symbolic information processing, ANN supports the expert systems and the combination holds very significant technological promise.

Uncertainties abound human endeavours and associated information and data are often vague. Conventional techniques are, therefore, less useful when it comes to modelling and newer tools such as ANN provide some hope. The theory of chaos has already pointed out that apparently random behaviour may, in fact, be deterministic. The notion of fuzziness that prevails has prompted the development of fuzzy models and tools based on fuzzy set theory and logic, and applications in engineering such as reliability modelling, critical-path optimization and fuzzy logic controls are gaining ground. Future developments will, therefore, be clearly dominated by these new emerging tools.

Seeking solutions through new seamless science

Chemical engineers are expected to solve problems. Sometimes one finds a tendency to skirt around these and try and model them. Consider the problem of fouling in polymer membrane devices. Endless number of papers have modelled this phenomenon. But just modelling fouling is being a 'part of a problem' and not a 'part of a solution'. We really need to think creatively of ways and means by which fouling can be prevented. The solutions may require again the use of most severe interdisciplinary science. For instance, now that we know how to use molecular information engineering, why not develop a membrane with a molecularly imprinted polymer? May be a smart self-cleaning surface using 2-D imprint with an attribute of an enzyme mimic might do the trick! Or why not use a piezoelectric polymer which enables generation of mild turbulence at the membrane surface for a continuous cleaning action? However, such solutions require CE scientists with borderless attitude, who will go beyond the traditional thinking and seek solutions from other fields, and this is the new generation of CE scientists we will have to create.

Let me speak from a personal experience on drag reduction. Over 2000 papers have appeared on drag reduction under turbulent-flow conditions in polymer solutions. But these papers have not helped us in resolving the issue of the basic limitation of these polymers, which arises due to their poor shear stability in turbulent fields. Many of us in the chemical engineering community have continued to model the degradation phenomenon. This modelling started in the seventies and continues even today in the nineties[73]. But have we solved the basic problem of *generating* shear-stable drag-reducing polymers? The only way we can solve the problem is by moving from the quest of 'new analysis' to new 'synthesis'. This new synthesis has to be again

based on an understanding of the molecular behaviour of these polymers under turbulent conditions, appreciation of the mechanochemistry of degradation and then finally designing some new synthetic strategies by using good contemporary chemistry.

We recently tried to make a modest effort towards this and achieved at least partial success. We used the knowledge of polymer chain breakage developed by polymer physicists[74]. Then we performed the mechanistic analysis of a perfectly 'zipped' assembly of a fully extended bead rod chain model of macromolecules in an elongational flow[75]. We showed that the enhanced stability of such zipped structures was a result of the distribution (near the chain centre) of the drag tension into the two parallel 'zipped' chains and reduction of the drag force due to the enhanced hydrodynamic shielding. This helped us in devising quantitative guidelines for the optimum design of shear-stable and effective drag reducers. But then we needed some ideas from chemistry to exploit this further. The final idea was to introduce some chemical moieties with interacting functionalities on the backbone of a polymer to bring in the association effect. We created a shear-stable drag reducer by using hydrogen bonding as a tool, which led to an interpolymer complex formed when dilute solutions of two polymers bearing complementary proton-accepting and proton-donating functional groups were mixed[76,77]. Zipped chains of the kind shown in Figure 12 were formed, with the 'zips' occurring due to hydrogen bonding. Such zipped molecules were far more shear-stable and also provided 2–6 times higher levels of drag reduction.

18

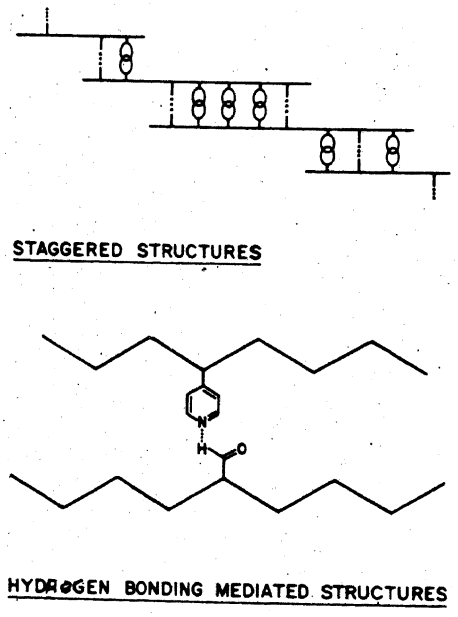


Figure 12. Design of shear-stable polymers by using hydrogen-bonding-mediated copolymeric structures, proton acceptor (containing 4-vinyl pyridine) and proton donor chains (containing methacrylic acid) create the staggered structures, which provide enhanced drag reduction[77].

The key point is that as CE scientists we had to take the clues from developments in polymer physics, perform an analysis based on a molecular model and finally shift to chemical synthesis of hydrogen-bonding-mediated systems to arrive at more effective systems. Remaining borderless or not shifting from 'analysis' to 'synthesis' would not have helped us to make progress. Of course, it is not that we claim that the problem of having shear-stable polymeric drag reducers is solved – but at least a modest beginning has been made.

Creating the seamless chemical engineering science: Some concerns

I have spoken on seamless chemical engineering science. I have referred to how the boundaries between different disciplines of sciences as well as engineering science are vanishing today. How do we make this boundaryless world work though? Let me now discuss some of my concerns in this respect.

First of all we all live in a world of very high specialization. This automatically implies that we are continuously erecting our own new boundaries, creating our own small clubs and generating our own special language and jargon all the time. However, boundarylessness implies interconnectedness. There itself lies the contradiction! How do we take care of these contradictions?

Another problem is that of dealing with adjacent sciences, where new mergers are continuously taking place and new territories are being continuously built. How do we deal with a moving target? What happens, for instance, when chemistry itself starts merging its boundaries with other disciplines, such as biology, in a profound way? Let me explain my worries by giving a specific example of the way chemists are trying to meet the challenge of generating and identifying biologically active compounds for pharmaceutical developments. Here, ideally, one would like to have access to an infinite source of compounds, encompassing the entire realm of structural and conformational space, that could be screened rapidly for individual molecules with desired biological properties. Chemists have now moved to combinatorial methods in biology and chemistry, which are powerful methods for generating molecular diversity (Figure 13). One approach, light-directed chemical synthesis, combines semiconductor-based photolithography technologies with solid-phase organic chemistry to synthesize large arrays of molecules with potential biological activity. A CE scientist is then confounded with the prospect of dealing with a phenomenon of rapidly emerging area, where boundaries between several adjacent disciplines have vanished. How does he cope with this challenge?

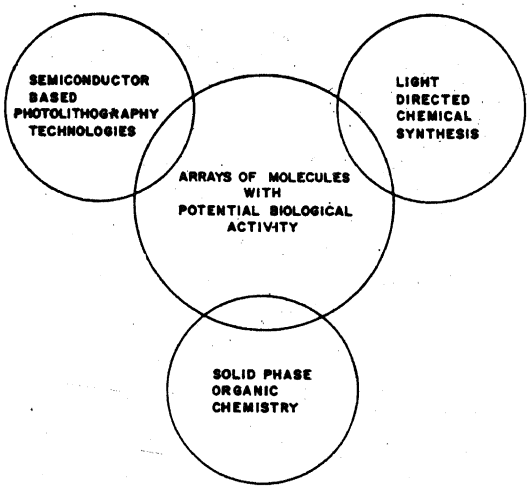


Figure 13. The principle of combinatorial chemistry, which combines diverse skills and technologies.

Then there is the problem of the language and the cultural barriers between different disciplines. Learning the language from another discipline and also teaching the language of CES to another discipline is not an easy job. Only those CE scientists who have dared to learn other skills and languages from adjacent disciplines will eventually make a mark. It also seems to me that interdisciplinary scientists will have to be great communicators too. They will have to constantly exchange information in a 'jargonless' way and also pass on their infectious enthusiasm by communicating well.

A CE scientist picks up tools and ideas generated in adjacent sciences and begins to publish papers in hardcore leading journals in chemical engineering. However, the elegance of the new tools or the sophistication of the new language may be mistaken for *novelty* and the problem solved may indeed be a trivial one. Many times a new tool from adjacent science, just because of its glamour value, may be applied to a trivial CE problem, which may not require that sophistication at all. This is misuse of tools and should be guarded against.

What are the forums for the presentation of results of research in interdisciplinary fields? Not many, I am afraid. More and more specialized clubs are mushrooming with little or no interaction between them. I saw someone make a suggestion that in a CES symposium dealing with an interface problem with other sciences, we should invite some scientists who may be able to give an alternative viewpoint. However, he was promptly turned down. Have we become too parochial? Will we be able to make progress if we continue this way?

We have another problem of peer recognition for those who work in interdisciplinary areas. As someone involved with the election of the Fellows of the National Academy of Science and National Academy of Engineering, I have found that engineers and scientists who are working on borders or trying to fuse these borders end up in difficulty in seeking recognition. Invariably, they fall between stools. Thus, somebody working in biomedical electronics does not get considered in either medical sciences or engineering sciences committees. Where does he go then? This is truly demotivating. The academies will have to recognize this problem and evolve special ways of solving these problems.

Many chemical engineering departments have recognized today the importance of having basic scientists with training and research experience from non-CE areas, be they in natural sciences or other branches of engineering. Thus, a chemical engineering department might employ an NMR specialist or a solid-state chemist. Sometimes the cultural and language barriers posed before these new entrants create great difficulties. A long time passes before they are accepted as part of the family. While speaking to some of them, it became clear to me that they also face difficulties on acceptance in academies or professional bodies. For instance, a physicist specializing in NMR in a chemical engineering department and applying NMR for solving problems of interest to chemical engineers does not get accepted as a Fellow of the Institute of Physics. At the same time the concerned Institute of Chemical Engineers also turns him down because it looks at his basic qualifications and finds that he is a non-chemical engineer! What a penalty to pay for working on borderline areas! Should not our professional bodies reexamine this and take new and bold approaches?

Then there is the question of how far does an engineering scientist go. There are some shining examples in our profession. Mark Davis from Virginia Polytechnic was educated as a traditional chemical engineer. He did his Ph D in kinetics and mathematical modelling. He wrote a book on numerical methods for chemical engineers. Then he contributed to the synthesis of zeolite A containing a transition metal. He exploited the use of ESR in the study of zeolites. Finally, he succeeded in synthesizing zeolite VPI-5, the first 18-ring opening crystal, which sparked off a renaissance in zeolite synthesis! A CE scientist is so well endowed due to his training that, if he dares to learn, he can conquer many territories. Extending this theme, I have reconsidered (Figure 14) the molecular heterogeneous hexagon conceptualized by van Santen[78]. It is clear that CE scientists will invade this hexagon in years to come with telling results!

With the boundaries between engineering science and science themselves merging, some basic conflicts are likely to arise. What is the role of an engineer and what is the role of a scientist? To my mind, a scientist 'explores', whereas an engineer 'exploits'. I looked for a

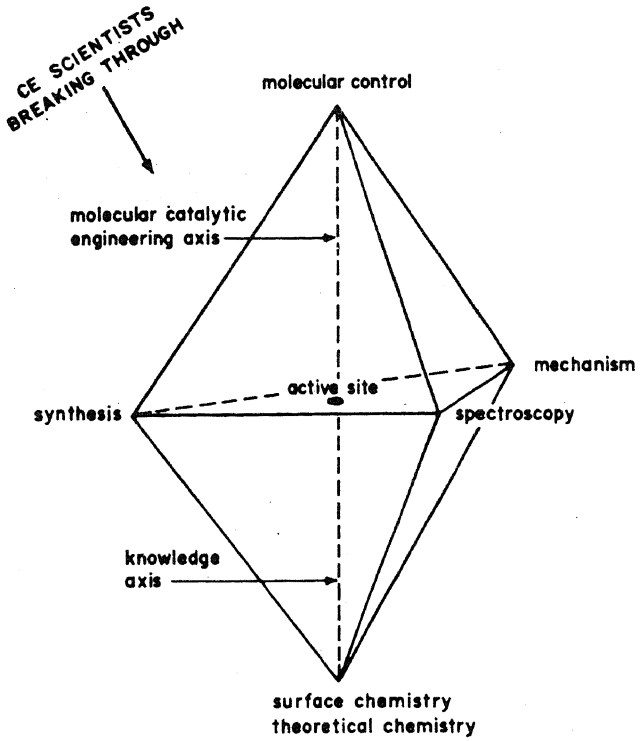


Figure 14. Molecular heterogeneous hexagon model due to van Santen[78]. Note the entry of CE scientists into this hexagon.

good definition of a scientist. I found one provided by Sir Peter Medawar. He says: 'A scientist is one, who weighs the earth and ascertains the temperature of the sun. He destroys matter and invents new forms of matter and one day he will invent new forms of life.' Then I looked for a definition of an engineer and found one given by Eric Laithwaite. He built on Sir Peter Medawar's definition and said: 'An engineer is one who uses the earth and tries to capture the sun's energy more effectively. He controls the rate of destruction of matter and tries to find alternative sources of energy and new materials. He invents new shapes of matter and strives to improve the quality of life in whatever form he finds it.' The basic goals of exploration and exploitation could not have been spelt out more clearly.

And that is precisely the problem that I am more concerned with. How does a CE scientist ensure that rather than getting possessed by the new tool that he has picked up and rather than spending all his research career on sharpening and refining this tool, does he bring benefit to the profession by using this tool in problem solving? Let me illustrate this point by going back to my early remarks about nonlinear dynamics and CES–physics interface.

Coupling between different nonlinear systems to bring about synchronization of states holds significant promise for future technology developments. Already, a few examples of the use of these concepts have emerged. Consider the spectacular success in enhancing the power of lasers by coupling them under chaotic conditions. In spite of the scores of papers on deterministic chaos, why is it that the CE scientists have not directed their attention to using this powerful technique to bring in similar spectacular results by, say, coupling chemical reactors? Here one is almost reminded of an inventor who endlessly refines his models but shies away from making a start for the fear that the real thing will not fly!

The challenges of the future will have to be met through both continuity and change. CES will extend itself, on the one hand, to the molecular focus of microscale processing and, on the other, to macroscale integration of processes and systems. Enormous flexibility and bewildering breadth have been the cornerstones of CES. However, the demand for enhanced flexibility in our approaches will be even higher in the coming years. The role of a CE scientist in this regard will be most crucial. He will have to be both very daring and innovative. In order to search for new tools and ideas, he may have to travel to the depths of the seas, climb new heights of the mountains, or he may have to go deep into the forests. If he does not dare to learn from other disciplines, then he will not be able to bring these jewels – these new ideas. The experience is that whenever this has happened, the results have been rewarding. Beckett *et al.*[79] borrowed the 'cold extrusion' idea from plastics and successfully applied it for the first time to milk chocolate, which is always melt-processed. This has opened up now endless possibilities of using plastics-processing techniques in chocolate processing. A CE scientist must not only dare to learn but he must also learn to dare.

Continuous improvements are fine – but discontinuities in knowledge create new technologies and that would not happen unless he dares. And most crucially, unless a CE scientist is himself 'boundaryless' and believes in seamless CES, all this will not happen.

Let us again remind ourselves that the word 'engineer' comes from the French word 'ingenieur', which literally means 'an ingenious one'. To my mind, our ingenuity of exploring the new tools of knowledge from adjacent disciplines is going to be the key to our success. Finally, I would like to repeat the inspiring inscription on the Lamme Medal of The Institute of Electrical and Electronics Engineers, USA, which simply says: 'The engineer views hopefully the hitherto unattainable.' I firmly believe that it is in this new seamless world of chemical engineering science that the CE scientists will not only 'view' the unattainable but also 'attain' the unattainable.

1. Danckwerts, P. V., *The Chem. Engr.*, July/August 1966, 155–159.
2. Lehn, J. M., *Science*, 1985, **227**, 849–856.
3. Amundson, N. R., *Frontiers in Chemical Engineering: Research Needs and Opportunities*, National Academy Press, Washington, DC, 1987.

4. Astarita, G., *Chem. Engng. Prog.*, 1990, 55–59.
5. Villermaux, J., *Chem. Engng. Sci.*, 1993, **48**, 2225–2235.
6. Field, R. J. and Burger, M., *Oscillations and Travelling Waves in Chemical Systems*, John Wiley, New York, 1985.
7. Ott, E., *Chaos in Dynamical Systems*, Cambridge University Press, Cambridge, 1993.
8. Doherty, M. F. and Ottino, J. M., *Chem. Engng. Sci.*, 1988, **43**, 139–183.
9. Razon, L. F. and Schmitz, R. A., *Chem. Engng. Sci.*, 1987, **42**, 1005–1047.
10. Schmitz, R. A., Graziani, K. R. and Hudson, J. L., *J. Chem. Phys.*, 1977, **67**, 3040–3044.
11. Jorgensen, D. V. and Aris, R., *Chem. Engng. Sci.*, 1983, **38**, 45–53.
12. Bandopadhyay, J. K., Ravi Kumar, V. and Kulkarni, B. D., *A. I. Ch. E. J.*, 1993, **39**, 908–912.
13. Lauterbach, J. and Rofermund, H. H., *Cat. Letts.*, 1994, **27**, 27–32.
14. Arce, P. and Ramakrishna, D., *A. I. Ch. E. J.*, 1991, **37**, 98–109.
15. Fox, R. O. and Villermaux, J., *Chem. Engng. Sci.*, 1990, **45**, 2857–2876.
16. Tritton, D. J. and Egdell, C., *Phys. Fluids A – Fluid Dynamics*, 1993, **5**, 503–505.
17. Allegre, C. J. and Turcotte, D. L., *Nature*, 1986, **323**, 123–127.
18. Aref, H., *J. Fluid Mech.*, 1984, **143**, 1–21.
19. Ottino, J. M., *The Kinematics of Mixing: Stretching, Chaos and Transport*, Cambridge University Press, Cambridge, 1989.
20. Ottino, J. M., *Annu. Rev. Fluid. Mech.*, 1990, **22**, 207–253.
21. Ottino, J. M., Muzzio, F. J., Tjahjadi, M., Franjione, J. G., Jana, S. C. and Kusch, H. A., *Science*, 1992, **257**, 754–760.
22. Howes, T., Mackley, M. R. and Roberts, E. P. L., *Chem. Engng. Sci.*, 1991, **46**, 1669–1677.
23. Sugawara, K., Hiroyuki, A., Sawada, Y. and Toko, K., *J. Phys. Soc. Japan*, 1993, **62**, 1143–1146.
24. Saido, Y., Matsuno, T., Toko, K. and Yamafuji, K., *Jpn. J. Appl. Phys.*, 1993, **32**, 1859–1866.
25. Lehn, J. M., *Angew. Chem. Int. Engl.*, 1990, **29**, 1304–1319.
26. Seebach, D., *Angew. Chem. Int. Ed. Engl.*, 1990, **29**, 1320–1367.
27. Mosbach, K., *TIBS*, 1994, **19**, 9–14.
28. Lightfoot, E. N., *Transport Phenomena and Living Systems: Biomedical Aspect of Momentum and Mass Transport*, John Wiley, New York, 1974.
29. Bailey, J. E. and Ollis, D. F., *Biochemical Engineering Fundamentals*, McGraw Hill, New York, 1977.
30. Lilly, M. D., *Chem. Engng. Sci.*, 1994, **49**, 151–159.
31. Frauenfelder, H. and Berg, H. C., *Physics Today.*, 1994, **No. 2**, 20–21.
32. Nielsen, J. and Villadsen, J., *Chem. Engng. Sci.*, 1992, **47**, 4225–4270.
33. Murlidharan Nair, T., Tambe, S. S. and Kulkarni, B. D., *FEBS Lett.*, 1994, **346**, 273–277.
34. Parton, R. F., Vankelecon, I. F. J., Casselman, M. J. A., Bezoukhanova, C. P., Vytternoeven, J. B. and Jocobs, P. A., *Nature*, 1994, **370**, 541–543.
35. Viatakis, G., Andersson, L. I., Muler, R. and Mosbach, K., *Nature*, 1993, **361**, 645–646.
36. Drexler, K. E., *Nanosystems Molecular Machinery, Manufacturing and Computation*, Wiley, New York, 1992.
37. Crandall, B. C. and Lewis, J., *Nanotechnology: Research and Perspectives*, MIT Press, Boston, 1992.
38. Whitesides, G. M., Mathias, J. P. and Seto, C. T., *Science*, 1991, **254**, 1312–1318.
39. Sara, M. and Sleytr, U. B., *J. Memb. Sci.*, 1987, **33**, 27–49.
40. Widawski, G., Rawiso, M. and Francois, B., *Nature*, 1994, **369**, 387–389.
41. Ley, S. V. and Low, C. M. R., *Ultrasound in Synthesis*. Springer, New York, 1989.
42. Prasad Naidu, D. V., Rajan, R., Kumar, R., Gandhi, K. S., Arakeri, V. H. and Chandrasekaran, S., *Chem. Engng. Sci.*, 1994, **49**, 877–888.
43. Martin, P. D. and Ward, L. D., *Trans. I. Chem. E.*, 1992, **70**, Part A, 296–303.
44. Muralidhara, H. S., *Chemtech.*, 1994, **No. 5**, 36–41.
45. Tsong, T. Y., *Biophys. J.*, 1991, **60**, 297–305.
46. Bowen, W. R., Kingdon, R. S. and Sabuni, H. A. M., *Memb. Sci.*, 1989, **40**, 219–229.
47. Asimov, I., *Chem. Engng. Progr.*, 1988, 43–47.
48. Osada, Y. and Ross-Murphy, S. B., *Sci. Am.*, 1993, May, 42–47.
49. Mashelkar, R. A., *J. Indian Inst. Sci.*, 1993, **73**, 193–214.
50. Wang, K. L., Burban, J. H. and Cussler, E. L., *Adv. Polym. Sci.*, 1993, **110**, 68–79.
51. Badiger, M. V., Kulkarni, M. G. and Mashelkar, R. A., *Chem. Engng. Sci.*, 1992, **47**, 3–9.
52. Kulkarni, M. G., Patil, S. S., Premanath, V. and Mashelkar, R. A., *Proc. Roy. Soc. (Lond.) A*, 1992, **439**, 397–407.
53. Osada, Y., Okuzaki, H. and Hori, H., *Nature*, 1992, **355**, 242–244.
54. Osada, Y., Okuzaki, H. and Gong, J. P., *TRIP*, 1994, **2**, 61–66.
55. Ozin, G., Kuperman, A. and Andreas, S., *Angew. Chem. Int. Ed. Engl.*, 1989, **28**, 359–376.
56. Larzlo, P., *Science*, 1987, **235**, 1473–1477.
57. Bridgwater, J., Broadbent, C. J. and Parker, D. J., *Trans. I. Chem. E.*, 1993, **71**, 675–681.
58. Broadbent, C. J., Bridgwater, J., Parker, D. J., Kenningley, S. T. and Knight, P., *Powder Technology*, 1993, **76**, 317–329.
59. Jasti, J. K. and Fogler, H. S., *A. I. Ch. E. J.*, 1992, **38**, 481–488.
60. Williams, R. A., Mann, R., Dickin, F. J., Ilyas, O. M., Ying, P., Edwards, R. B. and Rushton, A., *A. I. Ch. E. Symp.*, 1993, **89**, 8–15.
61. Gladden, L. F., *Trans. I. Chem E.*, 1993, **71**, 657–673.
62. Gladden, L. F., unpublished manuscript, 1994.
63. Devotta, I., Ambeskar, V. A., Mandhare, A. B. and Mashelkar, R. A., *Chem. Engng. Sci.*, 1994, **49**, 645–654.
64. Devotta, I., Badiger, M. V., Rajamohan, P. R., Ganapathy, S. and Mashelkar, R. A., *Chem. Engng. Sci.*, 1994, forwarded for publication.
65. Devotta, I., Premnath, V., Ganapathy, S., Badiger, M. V., Rajamohan, P. R. and Mashelkar, R. A., *Macromolecules*, 1994, **27**, 532–539.
66. Ganapathy, S., Badiger, M. V., Rajamohan, P. R. and Mashelkar, R. A., *Macromolecules*, 1989, **22**, 2023–2025.
67. Ganapathy, S., Rajamohan, P. R., Ray, S. S., Badiger, M. V., Mandhare, A. B. and Mashelkar, R. A., *Macromolecules*, 1994, **27**, 3432–3435.
68. Rajamohan, P. R., Badiger, M. V., Ganapathy, S. and Mashelkar, R. A., *Macromolecules*, 1991, **24**, 1423–1424.
69. Perkins, T. T., Smith, D. E. and Chu, S., *Science*, 1994, **264**, 819–825.
70. Kas, J., Strey, H. and Sackmann, E., *Nature*, 1994, **368**, 226–229.
71. *Bustamante*, C., Gurieri, S. and Smith, S. B., *TIBTECH*, 1993, **11**, 23–30.
72. Morris, A. J., Montague, G. A. and Willis, M. J., *Trans. I. Chem. E.*, 1994, **72**, Part A, 3–19.
73. Moussa, T. and Tiu, C., *Chem. Engng. Sci.*, 1994, **49**, 1681–1692.
74. Odell, J. A. and Keller, A., *J. Poly. Sci. Poly. Phys.*, 1986, **24B**, 1889–1916.
75. Agarwal, U. S. and Mashelkar, R. A., *J. Chem. Phys.*, 1994, **100**, 6055–6061.
76. Malik, S., Shintre, S. N. and Mashelkar, R. A., *Macromolecules*, 1993, **26**, 55–59.
77. Malik, S. and Mashelkar, R. A., *Chem. Engng. Sci.*, 1994, **50**, 105–116.
78. van Santen, R. A., *Chem. Engng. Sci.*, 1990, **45**, 2001–2011.
79. Beckett, S. T., Craig, M. A., Curney, R. J., Ingleby, B. S., Mackley, M. R. and Parson, T. C. L., 1994, *Trans. I. Chem E.*, 1994, **72**, 47–54.

REVIEW ARTICLE

The neurotransmitter receptor triad: Nicotinic acetylcholine receptor, γ-aminobutyric acid receptor (type A) and inhibitory glycine receptor

P. R. Satish and A. S. Balasubramanian

Department of Neurological Sciences, Christian Medical College Hospital, Vellore 632 004, India.

Many different neuroactive substances known as neurotransmitters, which are involved in signal transduction, bind to specific membrane receptors in the vertebrate central nervous system. Of particular interest are the similarities shared by the nicotinic acetylcholine receptor, γ-aminobutyric acid receptor and inhibitory glycine receptor with respect to their pentameric structure, their neurotransmitter binding pockets, and colocalization with gephyrin. All three receptors undergo phosphorylation catalysed by protein kinases A and C. The biochemical defects in relation to these receptors result in the clinical manifestation of certain human disorders.

RECENT studies have indicated that the agonist binding pockets of the excitatory nicotinic acetylcholine receptor (nAChR), the inhibitory γ-amino butyric acid receptor subtype A ($GABA_AR$) and the inhibitory neuronal glycine receptor (GlyR) are formed from conserved amino acid positions in the extracellular domains of the ligand-binding subunits. Alignment of the subunit sequences of these three receptors shows that a motif of aromatic side-chains equivalent to positions 159–161 of GlyR α1 subunit is highly conserved at the homologous positions of the ligand-binding $GABA_AR$ and nAChR proteins (Figure 1).

It has been demonstrated that the exchange of the aromatic residues at positions 159 and 161 of GlyR generates receptors that are preferentially gated by agonists with longer carbon chains, such as β-alanine, taurine or γ-amino butyric acid (GABA). Moreover, it was shown that the replacement of phenyl alanine at position 159 by a tyrosine residue was sufficient to convert the GlyR into GABA-responsive protein[1].

Another feature indicating a close relationship between nAChR, $GABA_AR$ and GlyR is the colocalization of gephyrin, a peripheral membrane protein, with GlyR-, nAChR- and $GABA_AR$-containing areas of brain. Further, experimental results indicate that determinants of subunit assembly are in part located at homologous positions in these three ligand-gated ion channel proteins[2–4].

The main objective of the following sections is to discuss the biochemistry of the ligand-gated ion channels (with emphasis laid on nAChR, $GABA_AR$ and GlyR), their modulation by phosphorylation and the defects underlying specific human disorders in relation to these receptors.

The biochemistry of ligand-gated ion channels

Ligand-gated ion channels constitute one of the two major classes of neurotransmitter receptors[5]. These are multimeric protein complexes; they serve to transduce the extracellular signals to the cell interior by acting, as their name suggests, as ion channels. The binding of their respective ligands induces the opening of these channels, which are then rendered permeable to specific ions. They are then said to convert the chemical signal (ligand binding) released from one cell into an electrical signal (hyperpolarization or depolarization) that propagates along the target cell membrane. Thus, they provide for rapid dialogue between cells of the central nervous system. (This class is exemplified by the excitatory nAChR, the inhibitory GlyR, the inhibitory $GABA_AR$, etc.).

To date six families of ligand-gated ion channel receptors that mediate information in brain and neuromuscular junction have been characterized at the level of their amino acid sequences[6,7]. They are: (i) the excitatory nAChRs of neuromuscular and neuronal origin, which conduct cations; (ii) the excitatory neuronal kainate-type glutamate-activated channels, which conduct cations;

GlyR	α1	154	M Q L E S	F G Y	T M N D L
GABA R A	α1	155	L K F G S	Y A Y	T R A E V
GABA R A	β1	152	L E I E S	Y G Y	T T D D I
nAChR	β1	144	M K L G T	W T Y	D G S V V

Figure 1. Alignment of partial GlyR, $GABA_AR$ and nAChR subunit sequences. The regions homologous to amino acids 154–166 of the GlyR1 subunit are shown; amino acids homologous to residues 159–161 of the GlyRα subunit are boxed. (Based on ref. 1)

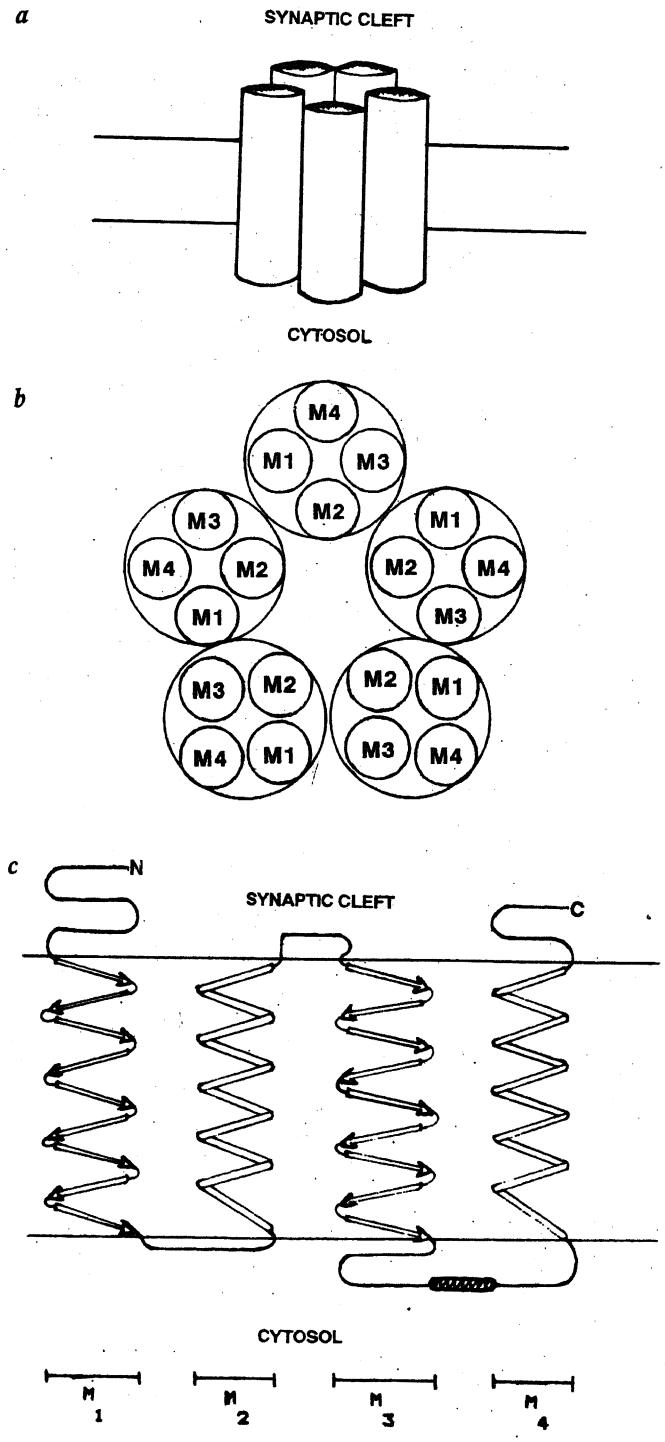


Figure 2. Schematic model of the topology of an idealized ligand-gated ion channel. *a*, Pentameric subunit structure viewed through an angle to the plane of the membrane. *b*, Schematic model representing the arrangement of the transmembrane domains, namely M_1, M_2, M_3 and M_4 (depicted as smaller circles) within each subunit (the larger circles) around the central ion channel (based on ref. 8). *c*, Membrane topology of an individual subunit containing transmembrane domains $M_1 - M_4$, viewed through the plane of the membrane. The hatched rectangle indicates the region of the subunit that contains the phosphorylation sites for various protein kinases. (Based on ref. 56)

(iii) the excitatory neuronal serotonin receptor (5-HT_3 receptor), which also conducts cations; (iv) the inhibitory neuronal GlyR, which conducts chloride; (v) the inhibitory neuronal $GABA_AR$, which also conducts chloride; and (vi) the sarcoplasmic reticulum ryanodine receptor, which conducts calcium.

These ligand-gated ion channel neurotransmitter receptors are pentameric complexes of subunits which are integral membrane proteins arranged around a central aqueous pore (Figure 2*a*, *b*)[8]. In a given receptor, these subunits reveal a high degree of sequence homology[9]. The sequence identity extends to other classes of receptors too. The hydropathy profiles displayed by the subunits of different families of these receptors are similar, suggesting that all possess a common transmembrane folding[9–14]. They are, therefore, said to form a superfamily of ligand-gated ion channels.

The minimal model for transmembrane organization of the subunits, which was proposed for Torpedo nAChR subunits[15,16], has been extended to all members of the superfamily[8]. It consists of

(a) a hydrophilic N-terminal domain oriented towards the synaptic cleft and carrying the glycosylated moieties,

(b) a compact hydrophobic region subdivided into three uncharged segments long enough to span the membranes and referred to as M_1, M_2 and M_3,

(c) a hydrophilic domain oriented towards the cytosol, which carries the consensus site for phosphorylation, and

(d) a short carboxy terminal hydrophobic transmembrane segment denoted as M_4.

In this minimal model both the N- and C-terminals are oriented towards the synaptic cleft (Figure 2*c*)[9]. However, this view has recently been challenged for glutamate receptor[17].

The extracellular domain of each of the receptor subunits forms a cylindrical vestibule which extends above the membrane plane. The agonist binding domains are located on the N-terminal domain and the cylindrical vestibule forms the ion channel, which lies along the axis of pseudosymmetry of the molecule[6]. Accordingly, the agonist site and the ion channel are topographically distinct. Hence, indirect or allosteric interactions account for the opening of the ion channels by their respective ligands[18].

Directed mutagenesis of the receptor subunits coupled with their transient expression in Xenopus oocytes has suggested that among the four hydrophobic putative membrane-spanning regions, namely M_1, M_2, M_3 and M_4, primarily M_2 and then M_1 and M_3 are most critical, and M_4 is the least required for ion conductance[5]. Therefore, each of the subunits contributes its M_2 region to line the inner wall of the ion channel[6,9]. Superimposed rings of charged residues bordering the M_2 regions of

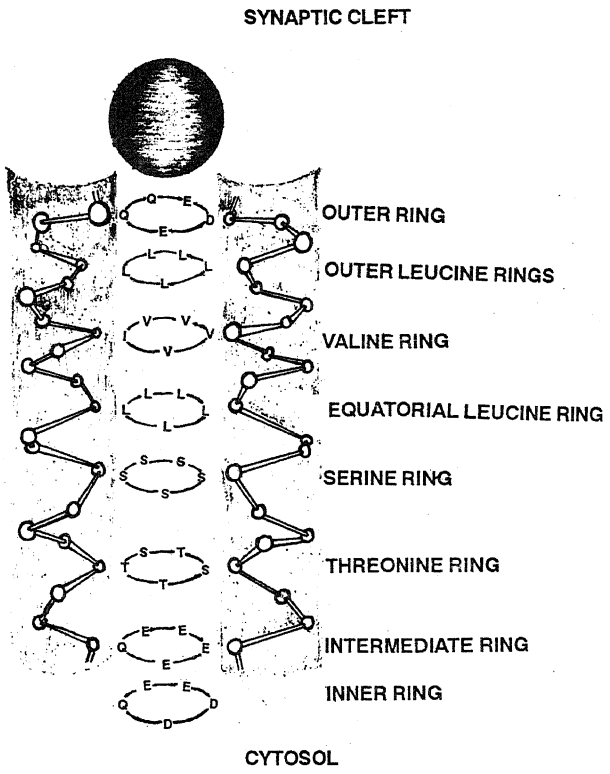


Figure 3. The arrangement of the M_2 segments of two subunits as transmembrane helices, wherein the superimposed rings of amino acids line the lumen of the ion channel. The channel is framed by rings of negatively charged amino acids at both the ends; these are called the inner ring and the intermediate ring at the cytosolic end, and the outer ring at the synaptic cleft end. At a low divalent ion concentration, removal of negative charges at these positions reduces both the inward and the outward currents. The other rings (from N- to C-terminal end) of the M_2 segments are: threonine ring, serine ring, equatorial leucine ring, valine ring and outer leucine ring. (Based on ref. 9)

the receptor have been shown to determine the conductivity of the ion channel (Figure 3). In all the subunits, charged amino acids are present at both ends of M_2.

In fact, high channel conductance of nAChR for cations is due in part to the negatively charged vestibule. The expected net negative charge within the vestibule is in excess of 10 negative charges, comprised of 31 negative and 21 positive groups. This excess negative charge density results in the repulsion of anions and concentration of cations in the channel mouth[19].

Significant in this regard is the fact that, $GABA_AR$ and GlyR, which conduct chloride ions (anions), have excess positive charge of about the same amount in their vestibular region. This suggests that charged vestibules may be a general mechanism of enhancing cation or anion selectivity[6].

The nAChR has a subunit stoichiometry of $\alpha_2\beta\gamma\delta$ in the neuromuscular junction[20], whereas it is just $\alpha_2\beta_3$ in the central nervous system and autonomic ganglia[21].

The subunit stoichiometry of GlyR which mediates synaptic inhibition in brain stem and spinal cord areas of the central nervous system (CNS) has been established[10] to be $\alpha_3\beta_2$, but which of the six α, four β, three γ, one δ and one ρ subunits comprise the $GABA_AR$ which mediates synaptic inhibition in other areas of the CNS is unclear as yet[22].

GlyR (ref. 23) and $GABA_AR$ (ref. 24) have been purified, and the amino acid sequence of their subunits and their functional expression studied[10–12]. Their subunits show sequence homology with each other to such an extent that the mutation of only two amino acids of the subunit of the GlyR renders it responsive to GABA[1,25]. These subunits have significant homology with the nAChR subunits also, further stressing the fact that all three receptors have a close relationship[1,7,9–12].

The combination of ion channels expressed by a neuron is responsible for its unique electrical properties like excitability and ability to propagate action potentials. The permeability of these ion channels to their respective ions, in addition to being regulated by membrane potential (in voltage-gated ion channels) and neurotransmitters (in ligand-gated ion channels), is also altered by protein phosphorylation[26].

Protein phosphorylation

Four decades since the first report implicating it in glycogen metabolism, protein phosphorylation has now come about to be acknowledged as the most common posttranslational modification involved in the regulation of protein functions.

Protein phosphorylation is regarded as the fundamental event in most cellular processes, and especially so in signal transduction[27,28]. Signals impinging on cells have their effects amplified and disseminated by a network of protein phosphorylation and dephosphorylation systems; the enzymes catalysing these reactions are further subject to control via cascades of phosphorylation and dephosphorylation reactions. Phosphorylation has also been shown to mediate many actions of the second messengers. Most extracellular signals including hormones, growth factors, neurotransmitters, electric potential and components of extracellular matrix alter the state of phosphorylation of intracellular proteins. These extracellular signals modify the activities of kinases and/or phosphatases either directly (e.g. receptors with kinase activity) or via cascades of enzymatic reactions (e.g. receptors–G-proteins–enzyme–second-messenger–protein-kinase) (Figure 4)[26]. Such cascades involving second messenger and protein kinase systems display a number of properties which allow amplification, integration, convergence and modulation of multiple and simultaneous internal and external signals.

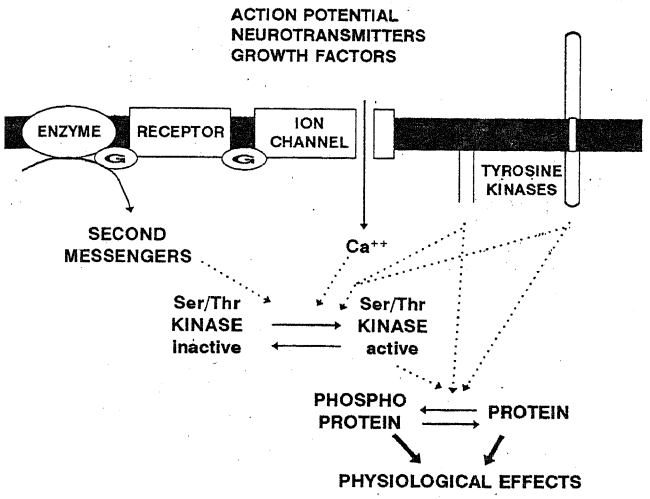


Figure 4. Schematic representation of some pathways by which extracellular signals alter protein phosphorylation in neurons. G in the figure indicates G (guanine nucleotide binding) protein.

Moreover, since the properties of many enzymatic components of signal transduction pathways are themselves altered by phosphorylation, the responsiveness of a cell to extracellular signal depends in part on the extent to which its proteins are phosphorylated. In other words, the cell keeps track of its recent history, encoded in its pattern of protein phosphorylation. Thus, protein phosphorylation allows not only the transduction of extracellular signals into the biochemical reactions inside the cells, but also the computation of several extracellular signals and their integration over time. It is, therefore, not surprising that most of the protein kinases and phosphatases are highly expressed in neurons, which are cells specialized in information processing[26].

Many proteins that are involved in synaptic function have been shown to be phosphorylated. These include proteins which play a role in neurotransmitter synthesis and release, voltage-gated ion channels, ligand-gated ion channels and G-protein-linked neurotransmitter receptors. The ion channels are a particularly appropriate target for the modulation of synaptic transmission by protein phosphorylation, since they are central to the process of signal transduction across the postsynaptic membrane[5,29].

Phosphorylation of the ligand-gated ion channels

Nicotinic acetylcholine receptor phosphorylation

Phosphorylation of nAChR has been extensively studied and documented[30,31]. One of the characteristics recorded for nAChR is its desensitization. In the continued presence of the agonist, the receptor rapidly fluctuates between conducting and nonconducting states for several hundred milliseconds and then enters a nonconducting, no-longer responsive desensitized state.

A comparison of the properties of the phosphorylated and nonphosphorylated nAChR shows that the rate of desensitization is enhanced in phosphorylated nAChR[32], whereas the channel conductance and the mean channel open time are not altered. Protein kinase activators like cyclic adenosine monophosphate (cAMP) analogues and phorbol esters are also recorded to enhance the rate of desensitization of nAChRs in different cell lines[33–36].

Tyrosine phosphorylation of nAChR is known to occur subsequent to innervation of muscle during development and is abolished by denervation. These findings establish the role of the neuron in regulating nAChR tyrosine phosphorylation[8].

Extracellular signals that are potential regulators of protein tyrosine phosphorylation have been investigated for their effect on the phosphorylation of nAChR. Agrin and basic fibroblast growth factor (bFGF) are the two well-documented examples. Agrin is known to induce tyrosine phosphorylation of nAChR in cultured myotubes[38] and the bFGF receptor is known to be a tyrosine kinase[37]. Whereas agrin mediates the nerve-induced aggregation of nAChR beneath the nerve terminals[8], bFGF bound to latex beads induces nAChR cluster formation in the Xenopus myoblast, at the bead myoblast contact point.

Drugs which inhibit tyrosine phosphorylation block both agrin-induced and bFGF-induced nAChR clustering[38]. Thus, tyrosine phosphorylation of nAChR most probably alters its interaction with the cytoskeleton in such a way as to cause the receptor immobilization, which results in accumulation of the receptor aggregates at the synaptic sites.

Phosphorylation of nAChR by calcium-diacylglycerol-regulated protein kinase or protein kinase C (PKC) in cultured chick myotubes results in the dispersal of nAChR aggregates and dephosphorylation promotes aggregation of the receptor. Diagonally opposite results are observed when phosphorylation due to cAMP-dependent protein kinase or protein kinase A (PKA) is activated[39].

All these facts suggest that phosphorylation, apart from enhancing the rate of desensitization of nAChR, also plays a vital role in its assembly and aggregation during synaptogenesis.

Recently, the autoregulation of phosphorylation of nAChR has been reported. This reveals a novel signal transduction pathway whereby the nAChR may regulate its own functional activity through phosphorylation[40].

GABA receptor phosphorylation

The evidence for direct modulation of GABA receptor function by phosphorylation is compelling. A basal-level

phosphorylation appears necessary to maintain $GABA_AR$ function[41]. Further, kinases activated in response to intracellular second messengers have been shown to regulate $GABA_AR$, resulting in a variety of functional effects, sometimes contradictory[42–44].

In different systems, for instance, PKA activation can potentiate, inhibit or have no effect on the amplitude of the GABA-elicited currents[45].

Phosphorylation due to PKC has been shown to decrease the current amplitude of some, but not all, GABA receptors.

These differential effects may be due to indirect effects of phosphorylation of other regulatory proteins; of equal importance is the fact that the functional properties of a given GABA receptor depend on the subunit composition. Different combinations of these subunits render the receptor responsive to different agonists and to a varied extent to the same agonist/antagonist[6,22].

Ethanol, man's most widely used and oldest psychoactive drug, acts by inducing the opening of $GABA_AR$ chloride ion channel. Recently, it has been reported that the alternatively spliced intracellular segment of the γ_2 subunit (namely, γ_2L, which has an eight amino acid insert), is required for the ethanol sensitivity in some GABA receptors. It is interesting to note that this intracellular segment contains a consensus PKC phosphorylation site[46]. Hence, it is possible that phosphorylation is involved in regulating receptor modulation at the ethanol-sensitive site.

Thus, in addition to their direct effects, intracellular phosphorylation events may also regulate GABA receptor more indirectly through effects on other modulatory sites.

Glycine receptor phosphorylation

GlyR can be phosphorylated both *in vitro* and *in vivo*, in response to PKA or PKC activators[47]. In Xenopus oocytes PKC activators decrease glycine-induced currents while PKA modulators increase the response to glycine. Although such data do not provide evidence for the actual phosphorylation of glycine receptor channels in the oocytes, the results obtained in intact spinal cord cells are consistent with a direct effect of phosphorylation in receptor function.

Given the variety of neurotransmitters, neuropeptides and growth factors known to regulate the intracellular activities of these two protein kinases, the fact that they modulate GlyR gives rise to intriguing possibilities, especially in the glycinergic neurotransmission in the spinal cord[48]. The dual mechanism of regulation by PKA and PKC modulators provides a sensitive instrument for the integrated modulation of the receptor by other extracellular messengers acting on the same neuron. This includes the previous, simultaneous or subsequent activation of a variety of G-protein-coupled receptors known to cause changes in cAMP or diacyl glycerol levels, which would further lead to changes in the subsequent neuronal response to glycine.

GlyR phosphorylation may prove to be a common mechanism by which the inhibitory action of glycine is regulated. This possibility is appealing, especially in case of pain transmission neurons. Two neurotransmitters, namely 5-hydroxytryptamine and noradrenaline, have important roles in the control of pain transmission at medulla and spinal levels. They are also known to alter the cAMP levels in target cells through their specific receptors. Thus, their observed effects may be mediated through the cAMP-dependent potentiation of inhibitory glycinergic transmission. Therefore, although further investigation is necessary, GlyR phosphorylation represents an attractive mechanism for the supraspinal regulation of some spinal cord processes. The integration of such cross-talking mechanisms of the neuronal membrane level accounts for transient and long-term changes in the efficacy of synaptic transmission, thus contributing to synaptic plasticity.

The defects underlying human disorders related to the neuroreceptors nAChR, $GABA_AR$ and GlyR

nAChR

Myasthenia gravis (MG) is a neuromuscular disorder characterized by weakness and fatigability of skeletal muscles. The underlying defect is a decrease in the number of available nAChRs at the neuromuscular junctions, due to an antibody-mediated autoimmune attack. Anti-nAChR antibodies cause accelerated destruction and the functional impairment of nAChR and failure of neuromuscular transmission. These antibodies act by three distinct mechanisms:

1. nAChR may be degraded at an accelerated rate by a mechanism involving crosslinking of the receptors.
2. The site in nAChR that normally binds acetyl choline may be blocked by the antibodies.
3. The postsynaptic muscle membrane may be damaged by the antibody in collaboration with the complement.

Thymus plays a role in the process of manifestation of autoimmunity. Muscle-like cells within the thymus, called myoid cells, bear nAChR-like protein on their surface. This acetylcholine-receptor-like protein is a glycoprotein, binds α-bungarotoxin irreversibly and has strong structural similarities with muscle nAChR. Its molecular weight, isoelectric point and amino acid composition are all very similar to those of nAChR found in mammalian muscle and the fish electroplax. This protein

most probably serves as the source of autoantigen and triggers the autoimmune reaction within the thymus gland[49–51].

$GABA_AR$

Epilepsies are a group of disorders characterized by chronic, recurrent paroxysmal changes in the neurological function caused by abnormalities in the electrical activity in the brain. Each episode of neurological dysfunction is called a seizure. Seizures may be convulsive when they are accompanied by motor manifestations or other changes in neurological function[52].

Current thinking awards GABAergic mechanisms a central role in the pathogenesis and treatment of epilepsy. The strongest evidence for a role of GABA in epilepsy is the fact that the pharmacological enhancement of GABAergic function produces antiepileptic effect. Several anticonvulsant drugs, including barbiturates and benzodiazepines, enhance GABA action at its primary sites, namely the GABA-benzodiazepine chloride ion channel complex[53].

GABA agonists are predominantly inhibitory and antiepileptic, whereas GABA antagonists produce seizures. Experimental agents that inhibit the breakdown of GABA are also predominantly antiepileptic.

Additional confirmation of the importance of GABA in epilepsy comes from analysis of neurotransmitter markers from animal models of acquired and genetic epilepsy. Pathological analysis of brain tissue from epileptic patients using positron emission tomography indicates either abnormalities in or the loss of GABA/benzodiazepine receptors[53].

GlyR

Reduced glycinergic inhibition caused by subconvulsive strychnine poisoning results in muscular hypertonia, heightened reflex excitability and an exaggerated response to sensory stimuli. These symptoms are similar to those of the autosomal dominant neurological disorder called startle disease, also known as hyperekplexia or kok's disease. This disorder is characterized by a marked muscular hypertonia in infancy, and an exaggerated startle response to unexpected sensory stimuli persistent in adulthood. Sudden jerking of limbs, or occasionally of trunk muscles, is observed; sudden noise or touch may cause the subject to jump or to fling an extremity.

More recently, point mutations in the gene encoding the subunit of the glycine receptor have been identified in startle disease pedigrees. These mutations map to arginine 271 located at the extracellular margin of the channel-lining M_2 region. In startle pedigrees the arginine 271 is replaced by an uncharged amino acid leucine or glutamine[54].

Site-directed mutagenesis of the GlyR cDNA coupled with their expression in mammalian 293 cells and their further studies using patch clamp technique have led to the discovery that the startle disease mutations cause no major alterations in ion permeation through GlyR, but rather dramatically reduce the receptor's agonist sensitivity. This finding, in addition to explaining the startle disease phenotype, identifies an unexpected structural determinant of the GlyR function.

This apart, analysis of other startle disease mutations may provide further important insights into the mechanisms of activation of the ligand-gated ion channel receptor superfamily.

Interestingly, startle response (abnormally enhanced startle response to acoustic stimuli) is also observed as one of the manifestations in the calcium calmodulin kinase II (CaMkII) mutant mice developed by Silva *et al.*[55]. The startle response is not observed in the wild-type mice, which have intact CaMKII. Whether there is any relationship between the startle response due to defective GlyR and that due to the absence of CaMKII, and if yes, the question as to what is its nature, has to be addressed in future.

1. Schmieden, V., Kuhse, J. and Betz, H., *Science*, 1993, **262**, 256–258.
2. Kirsch, J., Wolters, J., Triller, A. and Betz, H., *Nature*, 1993, **366**, 745–748.
3. Kirsch, J. and Betz, H., *Brain Res.*, 1993, **621**, 301–310.
4. Kuhse, J., Laube, B., Magalei, D. and Betz, H., *Neuron*, 1993, **11**, 1049–1056.
5. Huganir, R. L. and Greengard, P., *Trends Pharmacol. Sci.*, 1987, **8**, 472–477.
6. Stroud, R. M., Macarthy, P. M. and Shuster, M., *Biochemistry*, 1990, **29**, 11009–11023.
7. Marieq, A. V., Peterson, A. S., Brake, A. J., Myers, R. M. and Julius, D., *Science*, 1991, **254**, 432–437.
8. Swope, S. L., Moss, S. J., Blackstone, C. D. and Huganir, R. L., *FASEB. J.*, 1992, **6**, 2514–2523.
9. Thiery, A. D., Galji, J. L., Eisele, J. L., Bertand, S., Bertrand, D. and Changeux, J. P., *J. Memb. Biol.*, 1993, **136**, 97–112.
10. Betz, H., *Q. Rev. Biophys.*, 1992, **25**, 381–394.
11. Grenningloh, G., Rienitz, A., Schmitt, B., Methfessel, C., Zensen, M., Beyreuther, K., Gundelfinger, E. D. and Betz, H., *Nature*, 1987, **328**, 215–220.
12. Schofield, P. R., Darlison, M. G., Fujita, N., Burt, D. R., Stephenson, F. A., Rodriguez, H., Rhee, L. M., Ramachandran, Reale, V., Glencorse, T. A., Seeburg, P. H. and Barnard, E. A., *Nature*, 1987, **328**, 221–227.
13. Gregor, P., Mano, I., Maoz, I., Mckeown, M. and Teichberg, V. I., *Nature*, 1989, **342**, 689–692.
14. Hollmonn, M., O'shea-Greenfield, A., Rogers, S. W. and Heinemann, S., *Nature*, 1989, **342**, 643–648.
15. Claudio, T., Ballivet, M., Patrick, J. and Heinemann, S., *Proc. Natl. Acad. Sci. USA*, 1983, **80**, 1111–1115.
16. Noda, M., Takahashi, H., Tanabe, T., Toyosato, M., Kikyotani, S., Furutani, Y., Hirose, T., Takashima, H., Inayama, S., Miyata, T. and Numa, S., *Nature*, 1983, **302**, 528–532.
17. Seeberg, P. H., *Trends Neurosci.*, 1993, **16**, 359–365.
18. Herz, J. M., Johnson, D. A. and Taylor, P., *J. Biol. Chem.*, 1989, **264**, 12439–12448.
19. Imoto, K., Busch, C., Sakmann, B., Mishina, M., Konno, T., Nakai, J., Bujo, H., Mori, Y., Fukuda, K. and Numa, S., *Nature*, 1988, **335**, 645–648.

20. Cooper, E., Couturier, S. and Ballivet, M., *Nature*, 1991, **350**, 235–238.
21. Boulter, J., O'shea-Greenfield, A., Duvoisin, R. M., Connolly, J. G., Wada, E., Jensen, A., Gardner, P. D., Ballivet, M., Deneris, E. S., Mckinnon, D., Heinemann, S. and Patrick, J., *J. Biol. Chem.*, 1990, **265**, 4472–4482.
22. Raymond, L. A., Blackstone, C. D. and Huganir, R. L., *Trends Neurosci.*, 1993, **16**, 147–153.
23. Pfeiffer, F., Granham, D. and Betz. H., *J. Biol. Chem.*, 1982, **257**, 9389–9393.
24. Stauber, G. B., Ransom, R. W., Dilber, A. I. and Olsen, R. W., *Eur. J. Biochem.*, 1987, **167**, 125–133.
25. Grenningloh, G., Gundelfinger, E., Schmitt, B., Betz, H., Darlison, M. G., Barnard, E. A., Schofield, P. R. and Seeburg, P. H., *Nature*, 1987, **330**, 25–26.
26. Girault, J. A., *Neurochem. Int.*, 1993, **23**, 1–25.
27. Krebs, E. G. and Beavo, J. A., *Annu. Rev. Biochem.*, 1979, **48**, 923–959.
28. Edelman, A. M., Bluementhal, D. K. and Krebs, E. G., *Annu. Rev. Biochem.*, 1987, **56**, 567–613.
29. Hemmings, H. C., Nairn, A. C., McGuiness, T. L., Huganir, R. L. and Greengard, P., *FASEB J.*, 1989, **3**, 1583–1592.
30. Huganir, R. L. and Greengard, P., *Proc. Natl. Acad. Sci. USA*, 1983, **80**, 1130–1134.
31. Wagner, K., Edson, K., Heginbotham, L., Post, M., Huganir, R. L. and Czernik, A. J., *J. Biol. Chem.*, 1991, **266**, 23784–23789.
32. Hopfield, J. F., Tank, D. W., Greengard, P. and Huganir, R. L., *Nature*, 1988, **336**, 677–680.
33. Miles, K., Anthony, D. T., Rubin, L. L., Greengard, P. and Huganir, R. L., *Proc. Natl. Acad. Sci. USA*, 1987, **84**, 6591–6595.
34. Smith, M. M., Merlie, J. P. and Lawrence-Jr., J. C., *Proc. Natl. Acad. Sci. USA*, 1987, **84**, 6601–6605.
35. Middleton, P., Jaramillo, F. and Schuetze, S. M., *Proc. Natl. Acad. Sci. USA*, 1986, **83**, 4967–4971.
36. Mulle, C., Benoit, P., Pinset, C., Roa, M. and Changeux, J. P., *Proc. Natl. Acad. Sci. USA*, 1988, **85**, 5728–5732.
37. Ullrich, A. and Schlessinger, J., *Cell*, 1990, **61**, 203–212.
38. Nastuk, M. A. and Fallon, J. R., *Trends Neurosci.*, 1993, **16**, 72–76.
39. Green, W. N., Ross, A. F. and Claudio, T., *Proc. Natl. Acad. Sci. USA*, 1991, **88**, 854–858.
40. Miles, K., Audigier, S. S. M., Greengard, P. and Huganir, R. L., *J. Neurosci.*, 1994, **14**, 3271–3279.
41. Stelzer, A., Kay, A. R. and Wong, R. K. S., *Science*, 1988, **241**, 339–341.
42. Sigel, E. and Baur, R., *Proc. Natl. Acad. Sci. USA*, 1988, **86**, 6192–6196.
43. Browning, M. D., Bureau, M., Dudek, E. M. and Olsen, R. W., *Proc. Natl. Acad. Sci. USA*, 1990, **87**, 1315–1318.
44. Heuschneider, G. and Schwartz, R. D., *Proc. Natl. Acad. Sci. USA*, 1989, **86**, 2938–2942.
45. Moss, S. J., Smart, T. G., Blackstone, C. D. and Huganir, R. L., *Science*, 1992, **257**, 661–665.
46. Whiting, P., McKernan, R. M. and Iversen, L. L., *Proc. Natl. Acad. Sci. USA*, 1990, **87**, 9966–9970.
47. Vaello, M. L., Ruiz–Gomez, A., Lerma, J. and Mayor, Jr., F., *J. Biol. Chem.*, 1994, **269**, 2002–2008.
48. Song, Y. and Huang, L. M., *Nature*, 1990, **348**, 242–245.
49. Patrick, J. and Lindstrom, J. M., *Science*, 1973, **180**, 871–872.
50. Ahronov, A., Tarrab–Hazdai, R., Abrramsky, O. and Fuchs, S., *Proc. Natl. Acad. Sci. USA*, 1975, **72**, 1456–1459.
51. Drachman, D. B., in *Harrison's Principles of Internal Medicine* (eds Isselbacher, K. J., Brannwald, E., Wilson, J. D., Martin, J. B., Fauci, A. S. and Kasper, D. L.), McGraw Hill, New York, 1994, vol. 2, 13th edn, pp. 1–2489.
52. Dichter, M. A., in *Harrison's Principles of Internal Medicine*, (eds Isselbacher, K. J., Brannwald, E., Wilson, J. D., Martin, J. B., Fauci, A. S. and Casper, D. L.), McGraw Hill, New York, 1994, vol. 2, 13th edn, pp. 1–2489.
53. Johnston, M. V., in *Treatment of Epilepsy – Principles and Practice* (ed. Wyllie, E.), Lea & Febiger, Pennsylvania, 1993, pp. 1–1238.
54. Rajendra, S., Lynch, J. W., Pierce, K. D., French, C. R., Barry, P. H. and Schofield, P. R., *J. Biol. Chem.*, 1994, **269**, 18739–18742.
55. Silva, A. J., Stevens, C. F., Tonegawa, S. and Wang, Y., *Science*, 1992, **257**, 201–206.
56. Hucho, F., Gorne-Tschelnokow, U. and Strecker, A., *Trends Biochem. Sci.*, 1994, **19**, 383–387.

ACKNOWLEDGEMENT. P. R. S. is grateful to the Council of Scientific & Industrial Research, New Delhi, for award of a junior research fellowship. Work in the author's laboratory was supported by a grant from Department of Biotechnology, New Delhi.

Received 21 January 1995; accepted 17 May 1995

On the usefulness of quantum chemical bond order as a local reaction path indicator: The case study of a model carbonyl addition reaction†

D. K. Maity, D. Majumdar* and S. P. Bhattacharyya*

Chemistry Division, Bhabha Atomic Research Centre, Trombay, Bombay 400 085, India
*Department of Physical Chemistry, Indian Association for the Cultivation of Science, Jadavpur, Calcutta 700 032, India

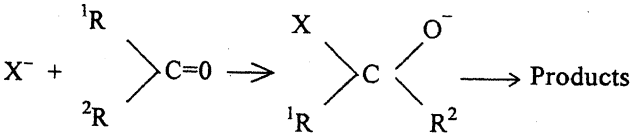
The idea of using quantum chemical bond order (BO) as a local reaction path indicator proposed earlier with reference to ABC ↔ BCA type of isomerization processes is explored further with reference to model carbonyl addition reactions, e.g. $H_2CO + H–Cl \rightarrow H_2CCl–OH$ and $H_2COH^+ + H–Cl \rightarrow ClCH_2OH + H^+$. In each case the reaction is viewed as a superposition of elementary chemical events, some representing bond making and some bond breaking processes, respectively. Altogether four bonds are involved in the model. The active bond order profile in each case is marked by a sharp inflexion signalling the onset of a critical process, viz. bond breaking or bond making. An extended bond energy bond order description of the addition of HCl to the >C=O moiety is suggested and tested numerically. The semilocal description is seen to model the approximate reaction path well and predict the active portion of the transition state structure fairly successfully. The local RP modelling is also seen to predict changeover in the mechanism of the reaction when conditions are altered by protonating the carbonyl group. The origin of the observed BO inflexions is sought to be correlated with Fermi correlation.

CHEMICAL reactions involve breaking and making of bonds. These processes take place gradually along a certain path, the so-called reaction path (RP). The usual indicator of this path is energy and the abstraction of RP in theory is a curving, parabolic potential energy channel stretching from the reactants to the product valley on the potential energy surface (PES). The top of the path defines the so-called saddle point, the activation barrier and the structure of the associated transition state. Energy is, however, a global parameter of the entire system while intuitive visualization of a chemical reaction in terms of breaking and making of specific bonds relies on a more 'local' description. An essential ingredient of this picture is the assumption that the bonds being broken or formed are the 'active' or 'important' bonds while the other bonds are more or less 'spectator like' or 'passive' and do not play any decisive role in shaping at least the gross or dominant features of the RP.

In the recent years a number of reports dealing with the existence of inflexion points in reactive valence or bond order (BO) profiles have appeared[1–5]. It has also been conjectured that such an inflexion point in BO profile may even coincide with or be close to the saddle point on the reaction path[4,5]. The present communication explores the utility of these concepts further with the specific case of a prototypical carbonyl addition reaction which can be described in terms of a four-bond process. It is also shown that by isolating the active bonds and making use of the appropriate bond order conservation conditions, an extended bond energy bond order (EBEBO) description of the active portion of the transition state geometry can be arrived at and the minimum energy reaction path (MERP) constructed. Although the discussion that follows concerns a reaction in particular, the observations are quite general or further generalizable. It enhances the scope of applicability of the local description to much more[6] complex systems.

Carbonyl addition reaction

Addition of nucleophiles to carbonyl centres is a very fundamental process in organochemical and biochemical reactions. Common examples include hydrolysis of esters and amides, aldol condensation and numerous reactions of organometallic reagents and ylides[7–9]. Analysis of experimental results on these reactions has led to the following B_{AC}^2 mechanism for these processes[10];



†Based on a lecture delivered by SPB in the 1993 mid-year meeting of the Indian Academy of Sciences, Bangalore.

Results of gas phase experiments have shown that esters undergo competitive proton transfer, eliminations and S_N2 reactions in favour of $B_{AC}2$ processes[11]. This $B_{AC}2$ mechanism has also been found to be operative for the gas phase additions to acyl halides[12]. Analysis of the kinetic results has led to the conclusion that the reactions in the gas phase feature a double-well potential surface with the tetrahedral species as a transition state and ion–dipole complexes as minima. Using crystallographic data for various nucleophilic groups interacting with different carbonyl centres, Burgi and Dunitz[13] were able to propose likely pathways for the approach of the nucleophile. Using SCF-LCAO-MO calculations[14] on the model addition reaction of hydride ion to formaldehyde, the same authors found no energy maxima in the potential energy profile. In a recent *ab initio* study on the addition of hydroxide ion to formaldehyde, Madura and Jorgensen[15], using a basis set containing diffuse functions, found a small local maxima in the energy profile of the reaction. Bayly and Grein[16] further confirmed that the reaction profiles for these addition reactions might be described adequately by adding diffuse functions to the basis set.

With this background in view, our purpose in the present communication has been to explore the possibility of (a) describing a simple carbonyl addition reaction quantitatively in terms of a rather *local* concept like variation of bond order, (b) constructing an *extended* bond energy bond order (semilocal) description of a process involving four active bonds, (c) using the local picture for predicting changeover in reaction mechanism, if any, when reaction conditions are altered, and (d) checking the validity of a *regional* or *local* hardness extremization principle proposed recently. The specific definition of bond order we choose to employ is due to Mayer[17]. The advantage of this quantity lies in its relative insensitivity to changes in the size of the basis set employed[18]. Irrespective of whether one is working in an *ab initio* or a semiempirical framework, the computed bond order index, which is a bond multiplicity index, turns out to be close to 1 for a single bond, 2 for a double bond, and so on. However, the minor method or basis dependence can be tackled by using renormalized BOs (see later). We have employed the standard AM1 method of Dewar and Thiel[19] with the hope of extending the present study, if useful, to reactions of different types involving much larger systems.

Method

The potential energy surface (PES) for the gas phase reactions studied in this paper has been computed by the standard AM1 Fletcher–Powell technique. For each point on the PES complete optimization of all geometrical parameters has been carried out. Bond orders have been calculated for each pair of atoms for all the relevant points on the RP. The generalized bond order (B_{XY}) between a pair of atoms (X and Y) is given by[17]

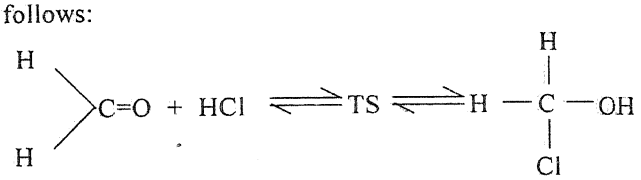
$$B_{XY} = \sum_{\mu \in X} \sum_{\nu \in Y} (\overline{D}S)_{\mu\nu}(\overline{D}S)_{\nu\mu} + \sum_{\mu \in X} \sum_{\nu \in Y} (D^{s}S)_{\mu\nu}(D^{s}S)_{\nu\mu}, \quad (1)$$

where the one-electron density matrix (ρ) in AO basis has been partitioned into $\overline{D}$ and D^s. $\overline{D}$ is the part of the one-electron density due to the core or the doubly occupied orbitals and D^s represents that due to singly occupied orbitals. At the AM1 or MNDO level of approximation, $S = 1$ and for closed shell systems $D^s = 0$, leading to corresponding simplifications in the expression for B_{XY} which now becomes identical with Wiberg's bond order[20].

Results and discussion

Addition of HCl to H_2CO in the gas phase: a local versus global description

The gas phase reaction between formaldehyde and hydrogen chloride may be schematically represented as follows:



We have explored two different pathways (schemes A and B in Figure 1) for the reaction, each passing through a separate transition state [TS(A)/TS(B)]. In scheme A, the HCl molecule has been assumed to approach the carbonyl group laterally, leading to a four-centered transition state [TS(A)] as in Figure 1. In scheme B, the HCl molecule is assumed to approach the nucleophilic centre in a plane perpendicular to the plane containing the >C=O group, leading to a transition state [TS(B)] as in Figure 1. In both the cases the reaction coordinate (RC) involves a predominant participation of the effective C..Cl and O..H distances. However, scheme B turns out to be inappropriate as it fails to lead to any well-defined TS or to product. In scheme A, on the other hand, we find that the reaction proceeds through a well-defined TS characterized by the structure [TS(A)] shown in Figure 1. This TS has been located by varying the >C..Cl and O..H distances (r_1 and r_2) in a preset grid while all other geometrical parameters for each pair of values of r_1 and r_2 have been fully optimized by the Fletcher–Powell technique.

The energy profile along the reaction path is depicted in Figure 2 *a*. The location of the TS corresponds to a value of the path coordinate (PC) = 0.76. At this point on the RP, $r_1 (= r_{C..Cl}) = 2.19$ Å, $r_2 (= r_{O..H}) = 1.34$ Å, $r_{C..O} = 1.30$ Å and $r_{H..Cl} = 1.63$ Å

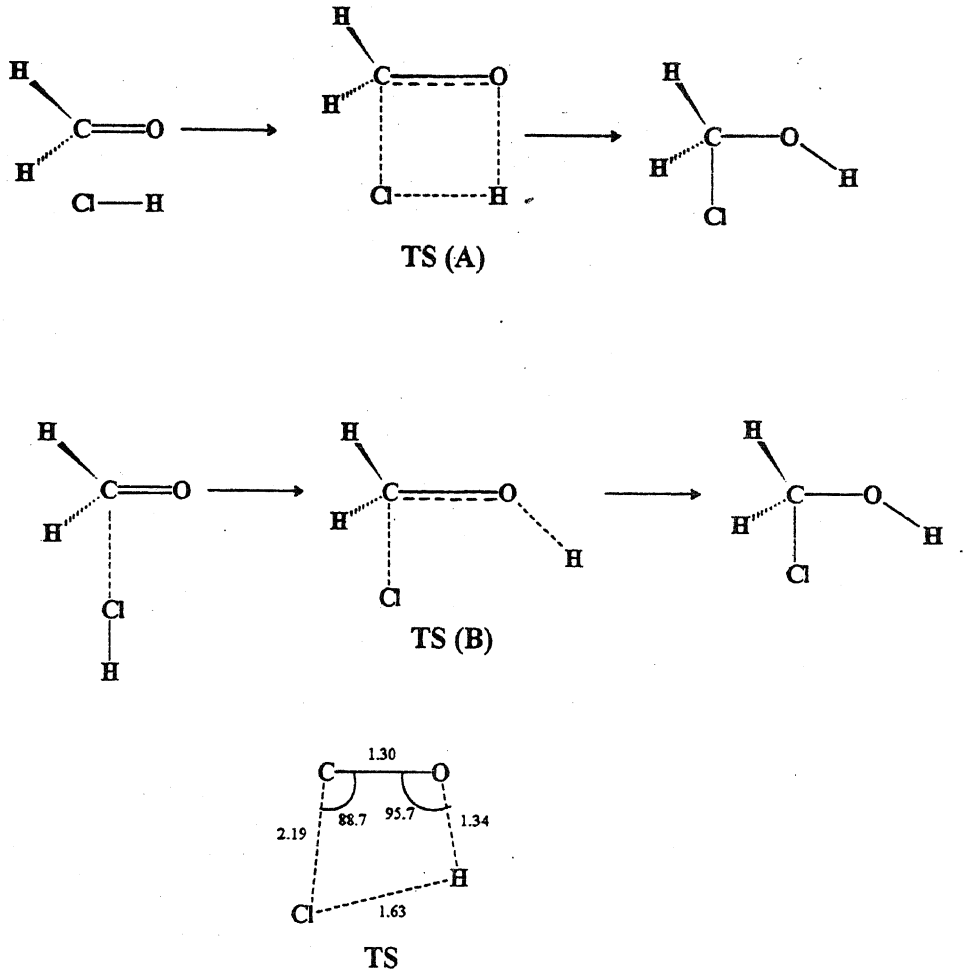
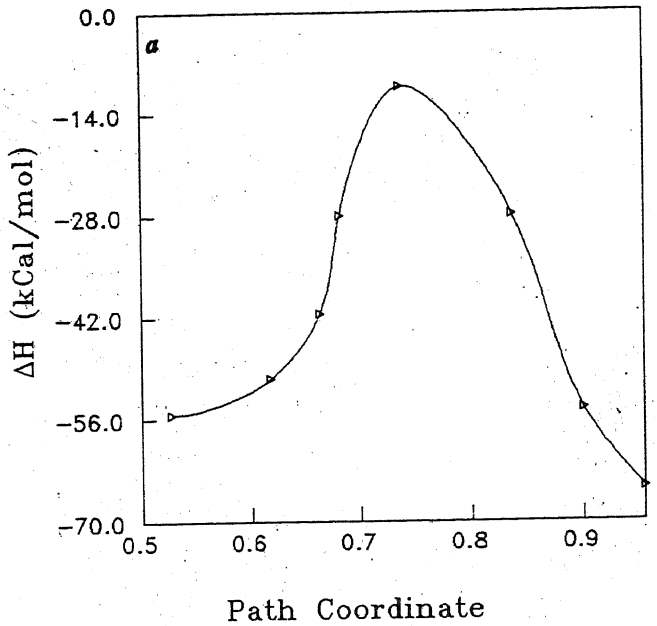


Figure 1. Reaction schemes (A and B) for the addition of HCl to H_2CO and optimized transition state geometry [TS(A)] for the lateral approach of HCl to the formaldehyde molecule.

(see Figure 1). Since the reaction scheme involves breaking of the H–Cl and p(C–O) bonds, respectively, and formation of a C . . Cl and an O..H bond, we designate them as 'active bonds'. A local model of the reaction should, therefore, be able to describe the process mainly in terms of quantities related to these four bonds only.

Figures 2 *b* and *c* display variations in H–Cl and O–H bond orders, respectively, as they change along RP, while Figure 3 *a* exhibits variations in the p bond order of the C–O bond as a function of the path coordinate. Variation in C–Cl bond order along the RP is also shown in the same figure (Figure 3 *b*). It is clearly seen that each of the H–Cl, C–Cl, C–O and O–H BO profiles reveals the occurrence of a definite and rather sharp point of inflexion. The C–Cl separation at the corresponding BO inflexion point is 2.19 Å, the H–Cl length is 1.63 Å , the C–O length is 1.30 Å and the O–H separation is 1.34 Å. If we compare these lengths with the bond length in the TS obtained by energy optimization route (i.e. by adopting a scheme of global RP analysis), we find that H–Cl, C–Cl, C–O and O–H bond lengths at the corresponding inflexion points match rather nicely with the corresponding bond lengths already found for the TS structure located on the given energy hypersurface.

It may be noted that the reaction in scheme A involves participation of four atoms (C, O, Cl and H). It would be of interest, therefore, to examine whether the total bonding capacity of the 'active atoms' is conserved along the MEP. We have, therefore, plotted the sum of the computed bond orders of the p(C–O), C–Cl, O–H and H–Cl bond (Figure 3 *c*) of the four relevant atoms along the RP. It is interesting to note that the so-called principle of bond order conservation (PBOC) advocated by Lendvay[2] appears to remain only approximately valid in this particular reaction in regions appreciably away from the critical region surrounding the saddle point. Around the saddle point, the active BO sum is not conserved. On the other hand, it passes through a minimum.



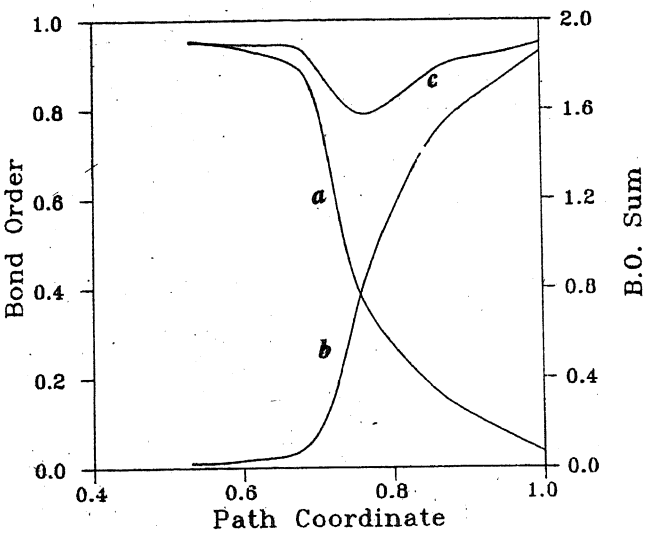


Figure 3. The BO profiles of: *a*, the Π (C–O) bond and *b*, the C–Cl bond along the reaction path for the gas phase addition of HCl to H_2CO. *c*, The extent of validity of the active BO sum conservation is shown for the carbonyl addition reaction.

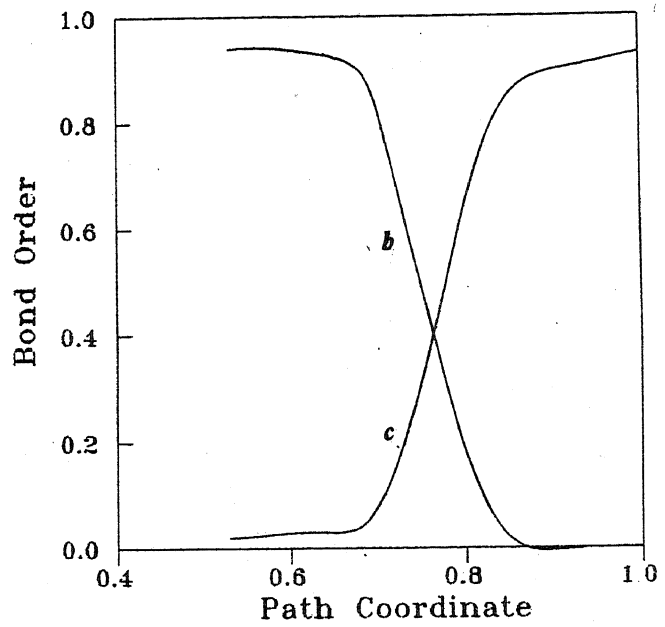


Figure 2. *a*, Energy profile in gas phase for formaldehyde +HCl reaction (scheme A): *b*, HCl bond order profile; *c*, O..H bond order profile along the RP.

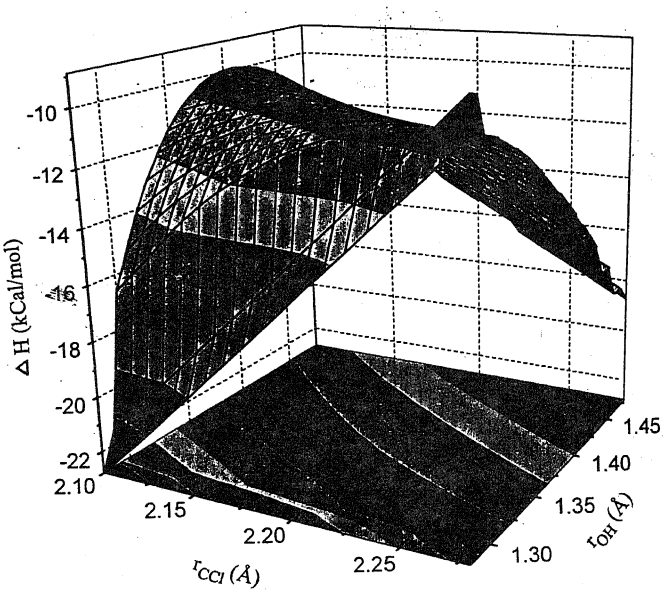


Figure 4. Energy contour as a function of O..H and C..Cl bond lengths for the formaldehyde +HCl reaction in the gas phase and the three-dimensional plot of the interaction energy. The transition state (×) on the reaction path (marked by the dotted line).

It has recently been shown that this behaviour of the active BO sum could amount to a local or regional softness or hardness extremization principle[21]. The saddle point, and the inflexion points in the BO profiles of the reaction under consideration occur almost at the same value of the path coordinate, viz PC = 0.76. The BO profiles of Figure 2 and 3, therefore, tend to suggest that the H–Cl, π(C–O), C–Cl and O–H bonds are each associated with a critical process occurring in the course of the reaction. They are then the active bonds. It would be worthwhile, therefore, to map the corresponding energy changes on the plane defined by the bond lengths of a pair of the active bonds, say C–Cl and O–H bonds. We have studied the reactive PES in three dimensions (two of these are the C–Cl (r_1) and O–H (r_2) bond lengths) and mapped its contour on the (r_1, r_2) plane. The energy map in two dimensions and the plot in three dimensions are displayed in Figure 4. The 'reaction path' has been marked on the contour map by a dotted line and the

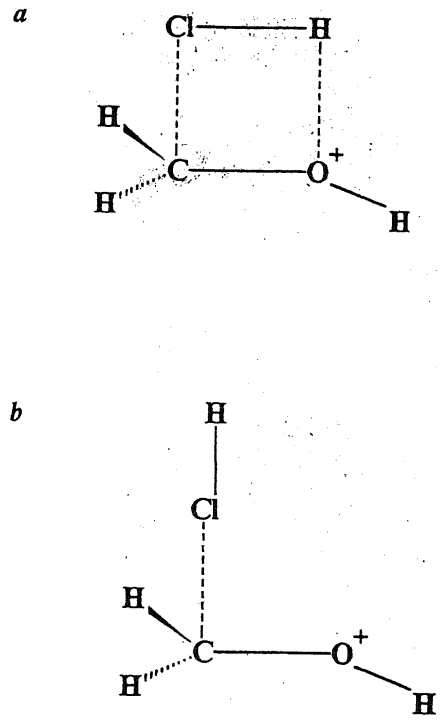


Figure 5. Schematic representation of: *a*, lateral and *b*, orthogonal approach of HCl to the protonated formaldehyde.

location of the TS by a cross (×) which corresponds to $r_1 = 2.19$ Å and $r_2 = 1.34$ Å. These values match well with those obtained from the points of inflexion on the BO profiles. The gas phase reaction is exoenergic by a small amount (5.34 kcal/mol) and the activation barrier turns out be ≈ 48.0 kcal/mol.

Gas phase addition of HCl to H_2COH^+: local description of mechanistic changeover

How would protonation of formaldehyde affect the course of addition of HCl to the carbonyl group? One can anticipate a drastic change in the mechanism since the addition of a proton would (i) severely distort the π electron distribution within the carbonyl moiety and (ii) create positive charge density on all atoms in the molecule. As expected, our calculations reveal the inappropriateness of the lateral approach, which fails to initiate the reaction (Figure 5 *a*). Initial search showed that the HCl molecule would prefer to approach the protonated >C=O group in a plane orthogonal to the plane containing the carbonyl group (Figure 5 *a*) with the Cl end pointing towards the carbonyl carbon atom. Figures 6 *a*

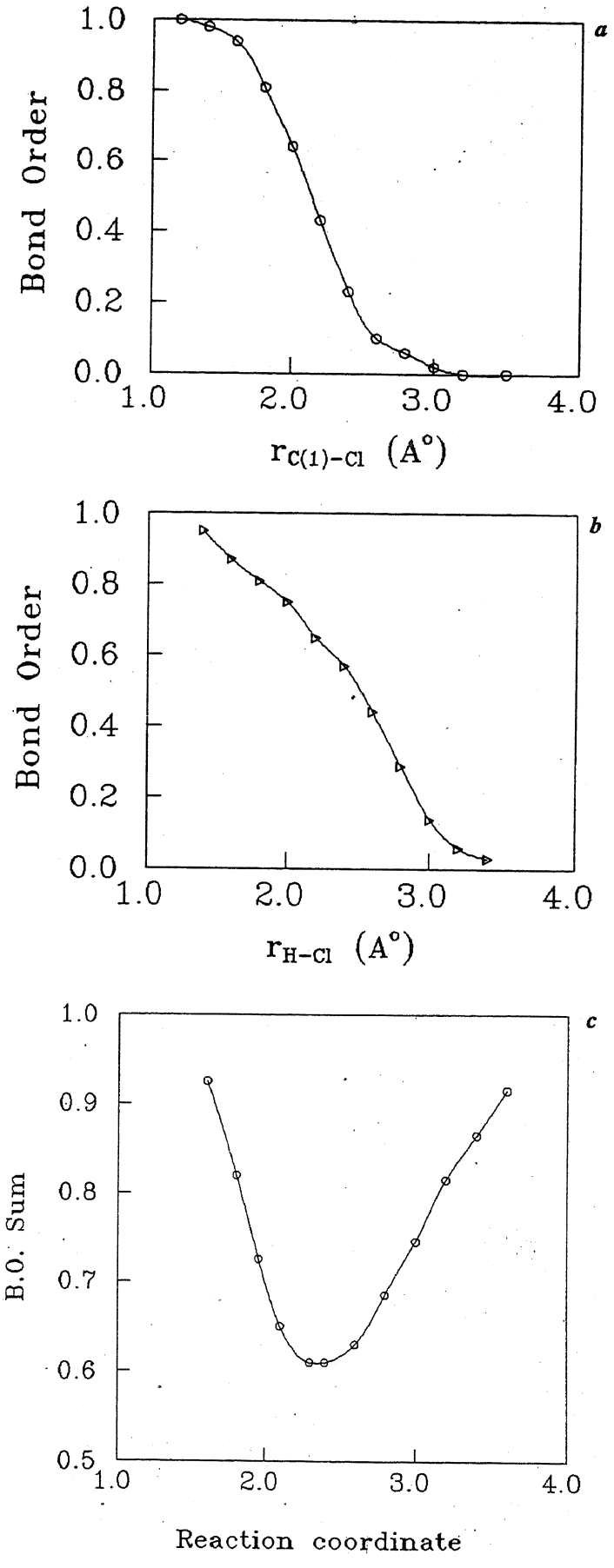


Figure 6. *a*, C–Cl bond order profile as a function of C..Cl bond distance; *b*, H–Cl BO profile as a function of H–Cl distance; *c*, sum of C–Cl and H–Cl BOs along the RP in the $HCl + H_2COH^+$ reaction.

and *b* present the H–Cl and C–Cl bond order profiles along the reaction path. These BO profiles, the active BO profiles, once again enable us to have a local description of the reaction in terms of elementary chemical events. It is interesting to note that the inflexion point in the BO profile of the H–Cl bond corresponds to an H–Cl distance of 2.75 Å when the C–Cl distance is about 2.25 Å. This C–Cl distance is much too large for any significant extent of C–Cl bond formation to have taken place. The inflexion point on the BO profile of the C–Cl bond itself, however, occurs only at a much shorter C–Cl length ($r_{C\text{–}Cl} \approx 2.08$ Å) indicating that the C–Cl bond formation takes place much later. In other words, *the breaking of the* H–Cl *bond does not synchronize with the formation of the* C–Cl *bond.* Obviously, this tends to suggest that the sum of H–Cl and C–Cl BOs is not conserved, a feature that is clearly seen to be true from Figure 6 *c*.

One can contemplate, therefore, that the mechanism of addition of HCl to H_2COH^+ could be fundamentally different from the mechanism established for the addition of HCl to unprotonated H_2CO. If the two asynchronous inflexion points, one rather early and the other late, observed in the H–Cl and C–Cl BO profiles, respectively, can be assumed to be associated with the physical process of bond breaking or making, one would be tempted to think that the addition of HCl to H_2COH^+ must be a two-step process. The H–Cl bond breaks first producing H^+ and Cl^-, the latter then attacking the carbon atom of the >C=O–H^+ group with the formation of chloromethanol. The entire process may be schematically represented as follows:

$$H_2CO + H^+ \rightarrow H_2COH^+,$$
$$HCl \rightarrow H^+ + Cl^-,$$
$$Cl^- + H_2COH^+ \rightarrow ClCH_2OH.$$

The inflexion that occurs in the C–Cl BO profile as Cl^- attacks H_2COH^+ is not very sharp, indicating that the associated energy profile also would be relatively featureless. Figure 7 *a* shows the energy profile for the first step when HCl approaches the carbonyl carbon atom of the protonated formaldehyde in an orthogonal plane. A small hump appears followed by a monotonic decrease in the total energy. The location of the hump corresponds to $r_{C\text{–}Cl} \approx 2.20$ Å and $r_{H\text{–}Cl} = 2.80$ Å, so that it may be taken to correspond to the breaking of the H–Cl bond, essentially corroborating what has already been revealed clearly by the bond order profiles. Figure 7 *b* represents energy profile for the second step of the reaction, in which chloride anion produced in the first step attacks the carbon atom of the H_2COH^+ moiety and formation of C–Cl bond of the product chloromethanol takes place. This step has been simulated by first determining the minimum-energy ground state geometry of H_2COH^+ and bringing in a chloride ion (Cl^-) from a large distance. The net charge of the

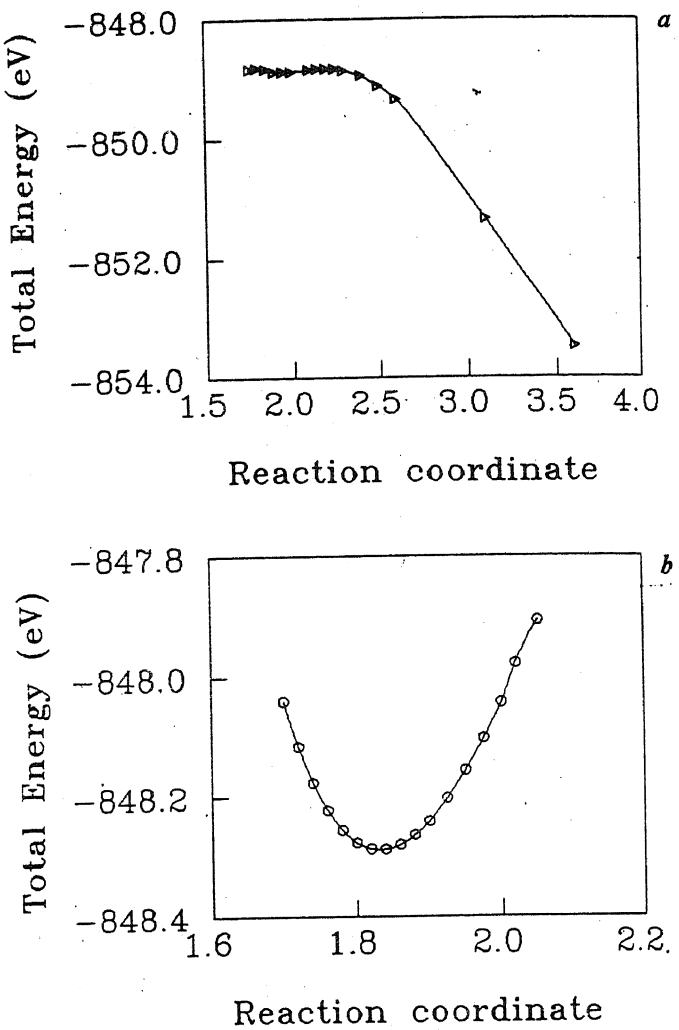


Figure 7. ***a***, Energy profile for the orthogonal approach of HCl to the protonated formaldehyde. The RC essentially represents the H–Cl bond length. ***b***, Energy profile for the orthogonal approach of Cl^- ion to the carbonyl carbon of protonated formaldehyde. The RC essentially represents the C–Cl bond length.

combined system has been kept zero and the entire system has been allowed to relax at each C–Cl distance. The energy decreases monotonically as chloromethanol is formed, without encountering any activation barrier. The first step, viz. the breaking up of HCl into H^+ and Cl^- must, therefore, be the rate-determining step in the addition of HCl^- to H_2COH^+. Relevant kinetic data are, however, unavailable to confirm this scenario experimentally. Similar behaviour has been noted for the addition of H–*X* (*X* = Cl, Br, CN, etc.) to the >C=O moiety, in general. The details will be presented elsewhere.

A BEBO description of the TS in the carbonyl addition reactions

One wonders at this point whether some kind of energy-based description of the carbonyl addition reaction is

feasible within the local model. In other words, how could we transcribe the local picture in terms of the BO profiles of active bonds into a semilocal energy-based picture? A possible recipe that immediately comes to mind is the BEBO model of Johnston and Parr[22, 23] which has been extensively used fairly successfully in studying atom transfer reactions[23–27] of the type $A + BC \rightarrow AB + C$. These reactions basically involve the active bond and the BEBO theory works well. In the present case, however, the basic model needs to be extended. Considering the four active bonds, viz. H–Cl, C–Cl, π(C–O) and O–H, and the corresponding bond orders, the sum of which is constrained to remain conserved along the reaction path, we have the following BEBO function for monitoring the energy changes during the reaction.

$$V = -D^0_{HCl} n^p_{HCl} - D^0_{CCl} n^q_{CCl} - D^0_{\pi(CO)} n^r_{\pi(CO)} - D^0_{OH} n^s_{OH}.$$

In short-hand notation, we may write (with $D^0_{HCl} = D^0_1$,

$D^0_{CCl} = D^0_2$, $D^0_{\pi(CO)} = D^0_3$ and $D^0_{OH} = D^0_4$),

$$V = -D^0_1 n_1^p - D^0_2 n_2^q - D^0_3 n_3^r - D^0_4 n_4^s \qquad (2)$$

where

$$n_1 + n_2 + n_3 + n_4 = 2; \qquad (3)$$

p, q, r and s are, parameters and D^0_{AB}, in general, is the energy of the A–B bond of unit order. Using the active BO sum conservation constraint (3) to eliminate n_4, we have

$$V = -D^0_1 n_1^p - D^0_2 n_2^q - D^0_3 n_3^r - D^0_4 \{2 - (n_1 + n_2 + n_3)\}^s \qquad (4)$$

As the TS, we have

$$\frac{\delta V}{\delta n_1} = \frac{\delta V}{\delta n_2} = \frac{\delta V}{\delta n_3} = 0. \qquad (5)$$

This leads us to a set of three coupled, non-linear equations:

$$\begin{aligned} -p\, D^0_1 n_1^{p-1} + s\, D^0_4\, \{2 - (n_1 + n_2 + n_3)\}^{s-1} &= 0, \\ -q\, D^0_2 n_2^{q-1} + s\, D^0_4\, \{2 - (n_1 + n_2 + n_3)\}^{s-1} &= 0, \\ -r\, D^0_3 n_3^{r-1} + s\, D^0_4\, \{2 - (n_1 + n_2 + n_3)\}^{s-1} &= 0, \end{aligned} \qquad (6)$$

and the BO conservation condition, which demands that

$$n_1 + n_2 + n_3 + n_4 = 2.$$

For $p = q = r = s = 2$, the equations in (6) are linear and can be solved together with $n_4 = 2 - (n_1 + n_2 + n_3)$. For other values of p, q, r, s equations (6) are to be solved iteratively. We have adopted a steepest descent (ascent) type of technique to solve the equations in (6) for $p, q, r, s \neq 2$. However, the TS BOs are found to be only weakly dependent on the p, q, r, s values. Our experience so far indicates that $p = q = r = s = 2$ is a good choice. The BO values obtained by solving equations (6) provide an approximate estimate of the active Bos in the TS. Using Pauling's bond length bond order relationship, $r_{AB} = r^0_{AB} - c^0_{AB} \ln n_{AB}$, the active bond lengths in the TS can also be calculated. The active bond lengths in the TS computed by our EBEBO recipe for (i) $p = q = r = s = 2$, (ii) $p = q = r = s = 3$, (iii) $p = q = r = s = 4$ are reported in Table 1 along with their corresponding values obtained from TS BOs predicted by the global energy-based analysis of the RP at the AM1 level of approximation (standard bond energy data have been used). We must mention here that Pauling's relation assumes that r^0_{AB} corresponds to the length of the A–B bond with BO = 1. However, frequently, the MO calculations predict equilibrium BOs of the A–B bonds that are different from 1. In such cases, one can use a renormalized Pauling bond length bond order relation:

$$r_{AB} = r^0_{AB} - c^0_{AB} \ln(n_{AB} / n^0_{AB}),$$

where n^0_{AB} is the computed bond order at the equilibrium internuclear separation (r^0_{AB}). The most important point to note is that it is the rate of change of BOs of different active bonds along the RP that dominantly determines where the critical point is going to be located. The absolute values of BOs are thus unimportant, rendering renormalized BOs useful and effective. Attempts have been made to improve the bond length bond order relation further by making c_{AB} dependent on the value of r^0_{AB}. Agmon's proposal[27] to take $c^0_{AB} = b_1^{AB} r^0_{AB} - b_2^{AB} r^2_{AB}$ invariably leads to further improvement. But in the present study, we have not used renormalized Pauling's relation or Agmons's improvement on it. When this is done, the agreement between extended BEBO and AM1 results could improve further. The values predicted by the EBEBO model do not fare too badly compared to values obtained from standard energy-based reaction path analysis. The V function of equation (2) can be used to generate energy surface or contour maps either in the BO space or in the internu-

Table 1. Comparison of the active bond lengths in the TS of the gas phase reaction $H_2CO + HCl \rightarrow ClCH_2OH$ computed by the EBEBO model and the AM1 method

Method		r_{HCl} (Å)	r_{CCl} (Å)	r_{OH} (Å)	r_{CO} (Å)
EBEBO	$p=q=r=s=2$	1.470	1.907	1.120	1.318
	$p=q=r=s=3$	1.470	1.930	1.162	1.321
	$p=q=r=s=4$	1.465	1.935	1.156	1.323
AM1		1.493	2.000	1.211	1.346

clear distance space. They are very much like the one described already and are, therefore, not reproduced here.

The origin of BO inflexions

We have already seen that only a few bonds are characterized by inflexions in their BO profiles along a given reaction path. That is, a reaction system appears to discriminate between two sets of BO variables, viz., the active and the inactive ones. The active BO variables are characterized by sharp inflexions along the reaction path. Why does it happen? It is clear that the BO inflexions occur over a rather short range. It will not be too speculative to assume, therefore, that forces responsible for the observed behaviour of the active BO profiles during a reaction must be changing rapidly over a rather short range. A clue to the nature of these forces may be hidden in the statistical interpretation of BO.

The chemical bond between a pair of atoms has been traditionally interpreted in terms of a build-up of electron density in the region (the bond region) between the two atoms. Customarily, therefore, the first-order density ρ_1 or quantities related to or obtainable from ρ_1 have been used for describing chemical bonds. A description based on ρ_1 cannot, however, reflect the 'electron pair' character of the most commonly encountered bonds. The second-order density matrix (ρ_2) is a more natural and logical candidate for analysing chemical bonds. The bond order index used by us is due to Mayer[17] and has been shown to be related to the normalization of the diatomic exchange part of the second-order reduced electron density (ρ_2). For a single determinant representable state, ρ_2 is factorizable and the bond order can then be expressed in terms of ρ_1. For such a state it has been shown[28] that

$$B_{AB} = -2\,\{\langle \hat{q}_A \hat{q}_B \rangle - \langle \hat{q}_A \rangle \langle \hat{q}_B \rangle\}, \tag{7}$$

where $\hat{q}_A$ and $\hat{q}_B$ are the LCAO atomic charge density operators for the atoms A and B, respectively, $\hat{q}_A$ and $\hat{q}_B$ are defined by

$$\hat{q}_A = \sum_{\mu \in A} \chi_\mu^+ \phi_\mu^-$$

$$\hat{q}_B = \sum_{\nu \in B} \chi_\nu^+ \phi_\nu^- \tag{8}$$

where χ_μ^+ and ϕ_μ^- are the 'creation' and the corresponding 'true-annihilation' operators corresponding to the spin orbital χ_μ on atom A. χ_μ^+ and ϕ_ν^- are defined accordingly for atom B. These operators satisfy the following relation:

$$\{\chi_\mu^+, \phi_\nu^-\} = \delta_{\mu\nu} \tag{9}$$

It can be shown that the right-hand side of equation (7) is related to the density–density correlation function C_{AB}, defined as

$$C_{AB} = \langle (\hat{q}_A - \langle \hat{q}_A \rangle)(\hat{q}_B - \langle \hat{q}_B \rangle) \rangle, \tag{10}$$

which immediately leads to the identification

$$B_{AB} = -2C_{AB}. \tag{11}$$

Now, $\hat{q}_A - \langle \hat{q}_A \rangle$ measures the fluctuations of electron density on atom A and $\hat{q}_B - \langle \hat{q}_B \rangle$ the same on atom B. C_{AB} in equation (10), therefore, represents the correlation between the fluctuations of the local electron densities on atoms A and B. The BO between two atoms A and B (B_{AB}) which we are monitoring is, therefore, twice C_{AB} (with sign reversed) and measures the correlation between these localized charge fluctuations. The force responsible for this correlation could be Fermi correlation only, as Coulomb correlation is absent in single determinant description of a state. At this stage a comment on the range of Fermi correlation would be in order. It has been shown by Cooper and Paunder[29] that the range of Fermi correlation appears to be somewhat longer than that of Coulomb correlation, contradicting observations made earlier by Boyd and Coulson[30] and also by Benesch and Smith[31]. However, the definition of Fermi correlation used by Cooper and Paunder is different. But even in the study made by Cooper and Paunder, the maximum and the most rapid change in Fermi correlation occurs over a rather short range and so Fermi correlation would abruptly decrease as a bond is stretched beyond the effective range of correlation as the system moves along a given RP. If two atoms are brought closer by such a motion, new correlations develop between charge fluctuations on the atoms involved and we may note an equally abrupt growth of a particular BO. It may be possible to propose another model to explain the observed behaviour of the active BO profiles based on orbital crossings and Walsh diagrams. But we have preferred to adhere to the explanation based on Fermi correlation in order to be consistent with the statistical interpretation. of BO.

Conclusions

Study of quantum chemical BO profiles may be quite helpful in visualizing the course of a chemical reaction. The BO inflexions provide us with a local view of a reaction involving the breaking and making of a number of bonds. We have been able to demonstrate the viability of an EBEBO description of a reaction involving up to a maximum of four active bonds. This brings quite a large variety of reactions under the ambit of our EBEBO model. They include (i) addition to >C=C<, (ii) ABC ↔ BCA isomerization, (iii) cyclopropyl$^+$ ↔ allyl$^+$rearrangement, (iv) reverse 1,3 dipole addition in tetra-aza cyclopentadienes, (v) intramolecular proton transfer in formamidene, (vi) 1,3 H shift, (vii) Cope rearrangement, etc. Many more reactions are being studied to assess the

overall applicability of the local model in studying chemical reactions.

1. Mayer, I., in *Modelling of Sructure and Properties of Molecules* (ed. Maksic, Z. B.), Ellis Haewood, Chichester, 1987, p. 45.
2. Lendvay, G., *J. Phys. Chem.* 1989, **93**, 4422–4429.
3. Ponec, R., *Collect. Czech. Chem. Commun.*, 1987, **52**, 2603–2612.
4. Maity, D. K. and Bhattacharyya, S. P., *J. Am. Chem. Soc.*, 1990, **112**, 3223–3225.
5. Maity, D.K., Das, K.K. and Bhattacharyya, S.P., *Proc. Indian Acad. Sci. (Chem. Sci.)*, 1989, **101**, 421–428; Majumdar, D., Maity, D. K. and Bhattacharyya, S. P., *Int. J. Quant. Chem.*, 1992, **43**, 567–572; Maity, D. K. Majumdar, D. and Bhattacharyya, S. P., *J. Mol. Struct. (Theochem.)*, 1992, **276**, 315–325.
6. Maity, D. K., Majumdar, D. and Bhattacharyya, S. P., *J. Mol. Struct. (Theochem.)*, 1995, **332**, 1–20.
7. Bender, M. L., *Chem. Rev.*, 1960, **60**, 53–113; Jencks, W. P., *Catalysis in Chemistry and Enzymology*, McGraw-Hill, New York, 1969.
8. Blow, D. M, *Acc. Chem. Res.*, 1976, **9**, 145-152.
9. March, J., *Advanced Organic Chemistry*, Wiley, New York, 1985.
10. Jencks, W.P., *Acc. Chem. Res.*, 1980, **13**, 161–169.
11. Faigle, I. F. G., Isolani, P. C. and Riveros, J. M., *J. Am. Chem. Soc.*, 1976, **98**, 2049–2052.
12. Asubiojo, O. I. and Brauman, J. I., *J. Am. Chem. Soc.*, 1979, **101**, 3715–3724.
13. Burgi, H. B., and Dunitz, J. D., *Acc. Chem. Res.*, 1983, **16**, 153–161.
14. Burgi, H. B., Lehn, J. M. and Wipff, G., *J. Am. Chem. Soc.*, 1974, **96**, 1956–1957.
15. Madura, J. D. and Jorgensen, W. L., *J. Am. Chem. Soc.*, 1986, **108**, 2517–2527.
16. Bayly, C. I. and Grein, F., *Can. J. Chem.*, 1989, **67**, 176–182.
17. Mayer, I., *Theo. Chim. Acta*, 1985, **67**, 315–322.
18. Mayer, I., Revesz, M. And Hargittai, I., *Acta Chim. Hung.*, 1983, **114**, 159–172; Mayer, I. and Revesz, M., *Inorg. Chim. Acta*, 1983, **77**, L205–L206.
19. Dewar, M. J. S. and Thiel, W., *J. Am. Chem. Soc.*, 1977, **99**, 4899–4907, 4907–4917.
20. Wiberg, K. B., *Tetrahedron*, 1966, **24**, 1083–1096.
21. Maity, D. K. and Bhattacharyya, S. P., *J. Phys. Chem.* (submitted 1995).
22. Johnston, H. S. and Parr, A. C., *J. Am. Chem. Soc.*, 1963, **85**, 2544–2551.
23. Johnston, H. S., *Gas Phase Reaction Rate Theory*, Ronald Press, New York, 1966.
24. Truhlar, D. G., *J. Am. Chem. Soc.*, 1972, **94**, 7584–7586.
25. Garret, B. C. and Truhlar, D. G., *J. Am. Chem. Soc.*, 1980, **102**, 2559–2570.
26. Garret, B. C., Truhlar, D. G. and Magnuson, A. W., *J. Chem. Phys.*, 1982, **76**, 2321–2331.
27. Agmon, N. and Levin, R. D., *J. Chem. Phys.*, 1979, **71**, 3034–3041.
28. De Giambiagi, M. S., Giambiagi, M. and Jorge, F.E., *Theo. Chim. Acta*, 1985, **64**, 337–341.
29. Cooper, I. L. and Paunder, C. N. M., *Int. J. Quant. Chem.*, 1980, **17**, 759–774.
30. Boyd, R. J. and Coulson, C. A., *J. Phys. B*, 1974, **7**, 1805–1816.
31. Benesch, R. and Smith (Jr), V. H., *J. Chem. Phys.*, 1971, **55**, 482–488.

ACKNOWLEDGEMENTS. We thank the CSIR, Govt. of India, New Delhi, for generous grant (No. 5(136)/88-EMR–II) and associateship to DM. DKM wishes to thank Dr J. P. Mittal and Dr K. V. S. Rama Rao, BARC, for constant encouragement.

Received 3 August 1994; revised accepted 21 June 1995

High-temperature tolerance of *Petunia* and *Nicotiana* pollen

G. U. Rao, K. R. Shivanna* and V. K. Sawhney§

Department of Botany, University of Delhi, Delhi 110 007, India
§Department of Biology, University of Saskatchewan, Saskatoon, Saskatchewan, S7N 5E2, Canada

High-temperature tolerance of the pollen of *Petunia hybrida* L. and *Nicotiana sylvestris* L. was investigated by treating dry pollen to temperatures of up to 75°C for 6–48 h and by studying their viability (by fluorochromatic reaction (FCR) test), vigour (by a semi-vivo method), and their ability to set fruits and seeds. In *Petunia*, temperatures of up to 60°C for 48 h did not affect pollen viability, vigour and their fruit- and seed-setting ability. A temperature of 75°C for 24 h reduced pollen viability and vigour, but fruit- and seed-setting ability existed. However, a 75°C exposure for 48 h proved lethal for *Petunia* pollen. In *Nicotiana*, pollen exposed to temperatures of up to 75°C for 6–12 h were able to set seed. With a longer exposure the majority of pollen were FCR-positive, but they were unable to set seed. These results show that pollen grains of *Petunia* and *Nicotiana* can withstand exposures of temperatures as high as 75°C and retain pollen function. This study also indicates that FCR test may not reflect true viability in pollen subjected to extreme stresses.

*For correspondence.

IN general, the activity of most metabolizing cells and tissues is best maintained within a moderate temperature range. Extreme low or high temperatures are detrimental to the activity or survival of most cells. Many plant cells respond to heat stress by the synthesis of heat shock proteins[1]. However, desiccated systems such as seeds can withstand a high degree of temperature range[2]. Pollen grains are also desiccated systems and show many physiological similarities to seeds[3]. The tolerance of pollen grains to low temperatures is well documented in studies on their storage under low and ultralow temperatures[4,5]. In contrast, much less is known of the high-temperature tolerance of pollen grains.

Earlier studies on high-temperature stress of pollen were restricted to examining the viability or *in vitro* germinability of treated pollen[6,7]. However, the effects of high temperatures on pollen function, i.e. pollen vigour and the ability of stressed pollen to induce fruit and seed set were not examined. Elsewhere, we have shown that *Brassica juncea* pollen exposed to 75°C for 24 h failed to germinate on the stigma and were unable to induce seed set[8].

In this paper we report on the viability, vigour and fruit- and seed-setting ability of *Petunia hybrida* and *Nicotiana sylvestris* pollen subjected to high temperatures of 45–75°C.

Materials and methods

Plants of *P. hybrida* L. grown under field conditions and of *N. sylvestris* L. grown under glasshouse conditions were used for investigations. Pollen grains from freshly dehisced anthers were uniformly spread on microslides, kept in dry Petri plates, and exposed to different temperatures: laboratory (22 ± 2°C), 45 ± 2°C, 60 ± 2C° and 75 ± 2°C, for various durations. After the treatment, pollen samples were tested for viability, vigour and their ability to set fruits and seeds.

Pollen viability was assessed by the fluorochromatic reaction (FCR) test as described by Heslop-Harrison and Heslop-Harrison[9]. Pollen grains were either used directly after the treatment or after prehydration of 30 min under high humidity by placing the samples in Petri plates lined with moist filter paper. Each treatment was repeated at least two times with two replicates each time. For each replicate at least 300 pollen grains were scored for the FCR test. Test of significance was performed by the group comparison test[10].

Pollen vigour was tested by the semi-vivo method[11,12]. Flower buds were emasculated before anthesis and pollinated on the day of anthesis with pollen exposed to different temperatures. Three hours after pollination, 1.0 cm of the style along with the stigma was cut with a sharp blade, and the cut end implanted in the pollen germination medium (sucrose 10% + boric acid 100 mg/l + calcium nitrate 300 mg/l + 0.8% agar) set in Petri plates. The cut ends of styles were observed under a stereomicroscope at regular intervals to record the time of pollen tube emergence and the number of tubes emerged. In one experiment, pollen tubes emerging from cut ends of styles were stained with DAPI (4′,6-diamidino-2-phenylindole), a DNA fluorochrome[13], to examine the division of the generative cell.

In vivo pollen germination and pollen tube growth was checked in pistils after 3–48 h of pollination. Pistils were fixed in FAA (formalin:glacial acetic acid: 70% ethanol; 5:5:90 v/v) for 24 h, cleared in 4 N NaOH overnight, and mounted in a drop of 0.005% decolourized aniline blue (Gurr) (in 0.05 M Na_2HPO_4 at pH 8.2). Pollen tube growth was examined in a Zeiss/Nikon fluorescent microscope.

Moisture loss in pollen grains was determined in 100 mg pollen samples. Pollen grains were placed in preweighed aluminium cups, and the samples maintained at different temperature conditions. After 24 and 48 h, pollen grains were weighed and the loss in weight was recorded as the loss of moisture content.

Pollen grains from all treatments were used for pollination of emasculated and bagged flowers to determine their ability to set fruit and seed. The flowers were rebagged after pollination. The mature fruits were harvested and the fruit size (diameter), weight of seeds per fruit, and the weight of 100 seeds of each treatment were recorded.

Results

Petunia hybrida

Moisture loss

There was no detectable loss of water in pollen samples maintained under laboratory conditions for 24 or 48 h, in comparison to fresh pollen (control). However, pollen exposed to 45 or 60°C temperature for 24 or 48 h resulted in approximately 5% water loss. An exposure of 75°C for 24 or 48 h caused an 8% water loss in pollen, and there was no further loss in pollen exposed for 48 h.

Pollen viability

Pollen viability was expressed as percentage FCR response. Exposures of 60°C for 24 h (Figure 1 *a*) and 45°C for 48 h (Figure 1 *b*) reduced the FCR values in comparison to fresh pollen (control) or lab conditions. However, 30 min prehydration of pollen restored the FCR score to the control level. Temperatures of 75°C for 24 h (Figure 1 *a*) and 60°C for 48 h (Figure 1 *b*) reduced significantly the percentage FCR response, and prehydration at these temperatures was not effective in restoring the FCR score. Pollen grains exposed to 75°C

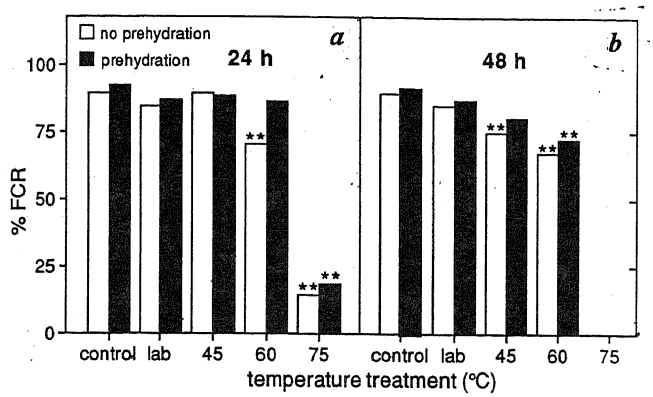


Figure 1. FCR response of *P. hybrida* pollen treated for 24 h (*a*) and 48 h (*b*) at different temperatures tested directly and after 30 min of prehydration. The values indicated by (**) are significantly different from the control (fresh pollen) at $P < 0.01$ level.

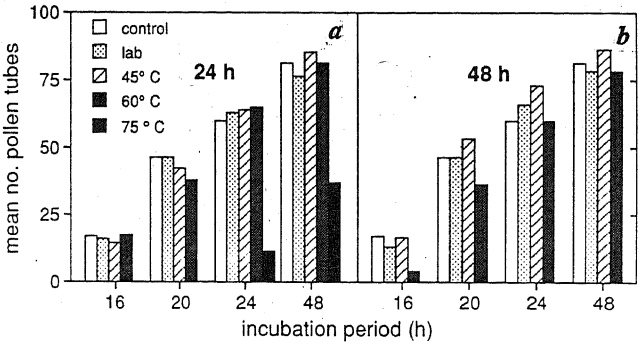


Figure 2. Mean number of pollen tubes ot *P. hybrida* emerging from semi-vivo implanted pistils pollinated with pollen exposed to different temperatures for 24 h (*a*) and 48 h (*b*), and incubated for 16–48 h. Pollen tube number was counted up to 100; it was not possible to count accurately beyond 100.

for 48 h failed to show the typical FCR fluorescence (Figure 1 *b*); they mostly showed dull fluorescence, although strong background fluorescence was observed in the medium.

Pollen vigour

The vigour of pollen grains, i.e. the time taken by pollen tube to travel through the style, was analysed using the semi-vivo method. The mean number of pollen tubes that emerged from the cut end of pistil varied depending on the temperature and duration of the treatment. In pollen exposed to 75°C for 24 h there was a marked delay in pollen tube emergence; no pollen tubes emerged for 20 h, there was some emergence by 24 h, and by 48 h the number of tubes emerged was less than one-half of that in the other treatments (Figure 2 *a*). In pollen grains exposed to 60°C for 48 h, there was reduction in the number of tubes emerging during the first 20 h; however, by 24 h the number was comparable to the control (Figure 2 *b*). There was no pollen tube emergence in pollen grains exposed to 75°C for 48 h (Figure 2 *b*).

Pollen tubes that emerged from styles were stained with DAPI to check the presence of sperm cells. In the control as well as in the treatments where pollen tubes emerged, two sperm cells were observed.

Pollen tube growth studies performed *in vivo* were in agreement with those of the semi-vivo method. The pistils pollinated with control pollen and with pollen treated for up to 60°C for 24 h showed high germination in 3 h, and a large number of pollen tubes reached the ovary in 24 h. However, pistils pollinated with pollen exposed to 75°C for 24 h showed very few germinating pollen in 3 h, and by 48 h only a few pollen tubes had reached the ovary. Pollen tubes of 75°C-treated pollen had a thick wall and the tube growth was not uniform compared to other treatments.

Fruit- and seed-setting ability

The fruit- and seed-setting ability of high-temperature-treated pollen showed that pollen grains treated for up to 60°C for 24 or 48 h were as effective (90% fruit set) as fresh pollen. Also, the size (diameter) of fruits was not affected in these treatments (Figure 3 *a*). Pollen exposed to 75°C for 24 h were also effective in inducing fruit set although fruit-set was observed in only 70% of pollinated pistils, and the fruit size was also reduced (Figure 3 *a*). The weight of seeds produced in a fruit was, interestingly, more in 45°C (both 24 and 48 h treatments) and in 60°C for 24 h, in comparison to the control (Figure 3 *b*), but it was reduced in the 75°C treatment (Figure 3 *b*). There was no difference in the weight of 100 seeds produced in most treatments. However, in 75°C (24 h) treatment, there was an increase in the weight of 100 seeds (Figure 3 *c*).

In one experiment 45 and 60°C treatments were extended to 5 days. There was no apparent effect on the fruit- and seed-setting ability of 45°C-treated pollen. However, the 60°C treatment reduced both the fruit size and the seed weight per fruit.

Nicotiana sylvestris

Preliminary experiments with *N. sylvestris* pollen showed that high temperatures of up to 60°C for 48 h had no adverse effect on pollen function. Therefore, studies on *Nicotiana* were largely confined to 75°C treatments for 6–48 h.

The pollen viability tests showed that over 50% of pollen exposed to 75°C for 24 h, were FCR-positive, in comparison to approximately 20% for *Petunia*. Exposures longer than 24 h were, however, lethal to *Nicotiana* pollen.

Pollen vigour tests by the semi-vivo method showed that in control more than 100 pollen tubes emerged in 20 h. In contrast, pollen exposed to 75°C for 12 h took

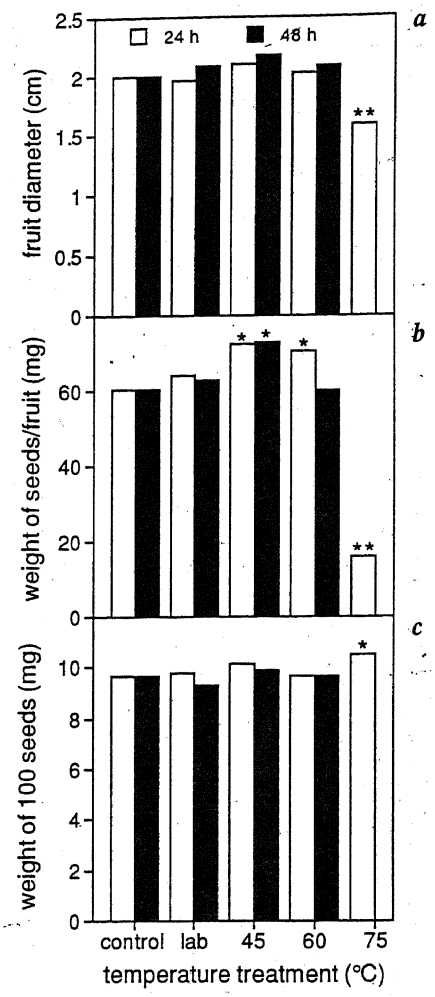


Figure 3. Fruit and seed characteristics obtained after pollination with pollen samples of *P. hybrida* exposed to different temperatures for 24 and 48 h: (*a*) fruit diameter; (*b*) average weight of seeds per fruit; and (*c*) average weight of 100 seeds. (*) and (**) are significant at $P < 0.05$ and $P < 0.01$ probability level, respectively, as compared to control (fresh pollen).

48–50 h for tube emergence and in those treated for 24 h there was no tube emergence even by 50 h. Fluorescence studies of pistils 50 h after pollination with pollen treated to 75°C for 24 h revealed germination of only 50–100 pollen grains on the stigma. Of these, only about 20–30 tubes entered the style and 5–10 tubes approached the cut ends of the style.

Fruit and seed sets were obtained with pollen treated at 75°C for 6 h and were comparable to the control (fresh pollen). However, pollen treated for 12 h yielded fruits in only 60% of pollinated flowers and the size of fruits and the number of seeds per fruit were markedly reduced (data not given). Pollen samples exposed to 75°C for 24 and 48 h failed to set fruit and seed.

Discussion

The results presented here have shown that pollen grains of *Petunia hybrida* and *Nicotiana sylvestris* are able to withstand a high temperature stress of up to 75°C, albeit for short periods, and retain not only the viability but also the ability to set fruits and seeds. However, the 75°C treatment reduces pollen vigour; treated pollen take much longer time to germinate and for the pollen tubes to reach the ovary. The reduction in the size of fruits and weight of seeds/fruit by 75°C (24 h) treatment in *Petunia* may be attributed to the fewer number of pollen tubes reaching the ovary before the ovules lose receptivity. However, there was a small increase in the weight of 100 seeds for this treatment. This may be explained on the basis of reduced competition among the fertilized ovules for nutrients.

FCR test is generally considered as the best available method to assess pollen viability (ability to effect fertilization)[14,15]. Our earlier studies[11] on pollen subjected to moderate stresses highlighted the limitation of the FCR test in assessing pollen vigour although it did reflect viability. Our studies on heat-treated pollen of *Brassica*[8] indicated that the extent of retention of fluorescence may give a better indication of pollen function than initial fluorescence. Present studies clearly show that the FCR test does not necessarily reflect true viability or vigour in pollen subjected to extreme stresses. Over 50% of *Nicotiana* pollen treated at 75°C for 24 h, were FCR-positive, but failed to effect fertilization. The pistils pollinated with such pollen showed germination of only 50–100 pollen on the stigma and growth of only 5–10 tubes in the style even after 50 h of pollination. Thus, the result of the FCR test on stressed pollen needs to be used with caution.

Previous studies on high-temperature tolerance of pollen were limited to a few taxa and were confined largely to *in vitro* germination tests. Pollen of lily, apple and rose showed[6] 50% germinability when exposed to 70°C for 4–8 h. A small proportion of *Eucalyptus rhodantha* pollen retained[7] *in vitro* germinability after exposure to 70°C for 24 h.

Our earlier studies on *B. juncea*[8] showed that exposure of pollen to 60°C for up to 24 h did not affect pollen viability and their ability to set fruits and seeds. However, pollen exposure to 75°C even for 4 h failed to germinate either *in vitro* or *in vivo*. The pollen of lily, apple and rose[6], *Eucalyptus*[7], and *Petunia* and *Nicotiana* (present studies) are all two-celled pollen systems, and seem to be more tolerant to high-temperature stress than the pollen of *B. juncea*[8], a three-celled system. Although high-temperature tolerance of no other three-celled pollen system, to our knowledge, has been investigated, this difference is in line with a range of other physiological differences between two- and three-celled pollen systems[16–19].

Heat tolerance of *E. rhodantha* pollen has been suggested to be an adaptation to the harsh semi-arid conditions prevailing in the natural habitat of the species (Western Australia), in which pollen grains are likely to remain exposed to prolonged heat and desiccation before pollination[7]. The authors suggested that similar heat tolerance is likely to be present in other species adapted for the same type of harsh environment as of *E. rhodantha*. The present investigation on *Petunia* and *Nicotiana*, and the earlier studies on lily, apple and rose[6], the species which grow under moderate or temperate conditions, have shown that high-temperature tolerance in all these species is comparable to that of *E rhodantha*. Thus, pollen grains in general and two-celled pollen in particular shown high-temperature tolerance, which is likely to be related to low moisture level present in most of the pollen systems. Small differences in the degree of tolerance between species may reflect the differences in the ability of the pollen protoplasts and membranes to withstand extreme desiccation associated with high temperatures.

Studies on pollen grains of some of the cereals such as wheat and maize also show that a low moisture level is an essential requirement for high-temperature tolerance. Pollen of cereals, unlike the pollen of a majority of taxa, are shed under more hydrated conditions[20]; such pollen lose their ability to effect fertilization even under moderately high temperature stress of 40°C for 4 h (ref. 21).

Extensive studies have been carried out on the tolerance of actively metabolizing organisms at supraoptimal temperatures[2,22]. Evidence indicates that the accumulation of metabolites such as proline and betain may provide thermal protection to several enzymes[23]. It is suggested that proline might stabilize protein configuration by maintaining hydration shells around the molecules. Pollen grains in general are rich in proline[4]. In *Petunia* pollen, free proline accounts for 2.6% of dry weight[24]. Zhang and Croes[25] reported that proline protected germination and metabolic function of *Lilium* pollen from unfavourably high and low temperatures. Further studies are, however, needed to understand the biochemical basis of thermotolerance in pollen grains. It has been shown that some heat shock proteins are expressed in developing seeds[26] and pollen grains[27]. These may also play a role in heat tolerance of desiccated pollen grains.

1. Vierling, E., *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, 1991, **42**,579–620.
2. Levitt, J., *Response of Plants to Environmental Stresses. Chilling, Freezing and High Temperature Stresses*. Vol. 1, Academic Press, New York, 1980.
3. Shivanna, K. R., Linskens, H. F. and Cresti, M., *Theor. Appl. Genet.*, 1991, **81**, 38–42.
4. Stanley, R. G. and Linskens, H. F., *Pollen: Biology, Biochemistry and Management*, Springer, Berlin, 1974.
5. Shivanna, K, R. and Johri, B. M., *The Angiosperm Pollen: Structure and Function*, Wiley Eastern, New Delhi, India, 1985.
6. Marcucci, M. C., Visser, T. and van Tuyl, J. M., *Euphytica*, 1982, **31**, 287–290.
7. Heslop-Harrison, J. and Heslop-Harrison, Y., *J. Cell Sci.*, 1985, **73**, 135–157.
8. Rao, G. U., Jain, A. and Shivanna, K. R., *Ann. Bot.*, 1992, **68**, 193–198.
9. Heslop-Harrison, J. and Heslop-Harrison, Y., *Stain Tech.*, 1970, **45**, 115–120.
10. Greig-Smith, P., *Quantitative Plant Ecology*, Butterworths, London, 1957.
11. Shivanna, K. R., Linskens, H. F. and Cresti, M., *Sex. Plant Reprod.*, 1991, **4**, 104–109.
12. Shivanna, K. R. and Rangaswamy, N. S., *Pollen Biology: A Laboratory Manual*, Springer, Berlin, 1992.
13. Coleman, A. W. and Goff, L. J., *Stain Tech.*, 1985, **60**, 145–154.
14. Heslop-Harrison, J., Heslop-Harrison, Y. and Shivanna, K. R., *Theor. Appl. Genet.*, 1984, **67**, 367–375.
15. Knox, R. E., Williams, E. G. and Dumas, C., *Plant Breed. Rev.*, 1986, **4**, 9–79.
16. Brewbaker, J. L., *J. Hered.*, 1957, **48**, 271–277.
17. Heslop-Harrison, J., *Int. Rev. Cytol.*, 1987, **107**, 1–78.
18. Hoekstra, F. A., *Planta*, 1979, **145**, 25–36.
19. Singh, S. and Sawhney, V. K., *Phytochemistry*, 1992, **31**, 4051–4053.
20. Heslop-Harrison, J. and Heslop-Harrison, Y., *Sex. Plant Reprod.*, 1992, **5**, 247–255.
21. Dupuis, I., *Int. Rev. Cytol.*, 1992, **140**, 391–405.
22. Sullivan, C. Y., Jordan, W. R., Blum, A. and Traore, M., in *Proceedings of the International Congress of Plant Physiology*, Vol. 2 (Sinha, S. K., Sane, P. V., Bhargava, S. C. and Agrawal, P. K, eds) Society of Plant Physiology and Biochemistry, New Delhi, 1990, pp. 916–922.
23. Paleg, L. J., Doughlas, A. J., van Daal, A. and Keech, D. B., *Aust. J. Plant Physiol.*, 1981, **8**, 107–114.
24. Linskens, H. F. and Schrauwen, J. A. M., *Acta Bot. Neer.*, 1969, **18**, 605–614.
25. Zhang, H. Q. and Croes, A. G., *Plant Cell Environ.*, 1983, **6**, 471–476.
26. Hernandez, L. D. and Vierling, E., *Plant Physiol.*, 1993, **101**, 1209–1216.
27. Van Herpen, M. M. A., Reijnen, W. H., Schrauwen, J. A. M., de Groot, P. F. M. and Wullems, G. J., *Sexual Reproduction in Higher Plants* (Cresti, M., Gori, P. and Pacini, E., eds), Springer, Berling, 1988, pp. 277–282.

ACKNOWLEDGEMENTS. We thank Mr Dennis Dyck for assistance in the preparation of figures. G.U.R. was supported by a Senior Research Fellowship from the Council of Scientific and Industrial Research, New Delhi. This study was supported in part by a Natural Sciences and Engineering Research Council of Canada Operating Grant to V.K.S.

Received 14 February 1995; revised accepted 25 May 1995

Enhanced giant magnetoresistance in magnetic/nonmagnetic $La_{1-x}MnO_{3-\delta}$ oxide multilayers

S. Sundar Manoharan, K. M. Satyalakshmi*, V. Prasad†, S. V. Subramanyam† and M. S. Hegde‡

Solid State and Structural Chemistry Unit, *Department of Metallurgy, †Department of Physics, Indian Institute of Science, Bangalore 560 012, India

Over 80% negative giant magnetoresistance (GMR) is observed in the self doped $La_{0.7}Mn_{3-\delta}$ oxide thin film in a magnetic field of 6 Tesla at 225 K. $La_{0.7}Mn_{3-\delta}$ (magnetic)/$La_{0.85}MnO_3$ (nonmagnetic) oxide multilayers showed a higher GMR effect up to 89% at 225 K compared to the single $La_{0.7}Mn_{3-\delta}$ oxide films. The results demonstrate that the large GMR observed in the doped manganates is inherent to the oxide material whereas the enhanced GMR to the extent of 6–10% in the multilayers compared to the parent magnetic film is due to the coupling of magnetoresistance in the parallel magnetic oxide layers separated by nonmagnetic oxide layers as in $(Fe/Cr)_n$ multilayer metal film.

GIANT magnetoresistance (GMR) is defined by the ratio $(R_H - R_0)/R_0$, where R_H and R_0 are the resistances in presence and absence of a magnetic field. GMR effect was first observed in the single crystals of $La_{0.62}Pb_{0.38}MnO_3$[1]. Over 50% negative GMR was reported by Baibich *et al.*[2] in $(Fe/Cr)_n$ magnetic/nonmagnetic metal multilayers. Presently, there is a tremendous interest in the study of GMR effect because, large variation in the resistance in the presence and absence of a magnetic field can have direct application as the 'readhead' for magnetic recording systems. Observation of over 50% GMR in the epitaxial thin films of $La_{0.67}Ba_{0.33}MnO_3$ oxide thin films[3] at 300 K triggered the interest on this subject. $La_{1-x}M_xMnO_3$ (M = Ca, Sr, Pb)[4–6] thin films showed large GMR effects.

$LaMnO_3$ is an antiferromagnetic oxide possessing orthorhombic structure. It is possible to synthesize lanthanum deficient $La_{1-x}MnO_3$ with 18–20% Mn ions in the 4+ state. These compounds crystallize in a rhombohedral structure and exhibit a paramagnetic to ferromagnetic transition, coupled with insulator to metal transition, when samples are cooled below 300 K. We have recently observed over 80% GMR in 5% Ag-doped $La_{0.7}MnO_{3-\delta}$ thin films at 220 K.

Although large GMR effect is observed in the doped manganates, it is not yet clear whether the mechanism of GMR operating in these oxide films is the same as in the $(Fe/Cr)_n$ multilayer metal films. In order to gain further insight into the mechanism of GMR we considered it worthwhile to fabricate $La_{0.7}MnO_{3-\delta}$ (magnetic)/$La_{0.85}MnO_{3-\delta}$ (nonmagnetic) oxide multilayers since they are structurally compatible. We report, in this paper, the first fabrication of magnetic/nonmagnetic oxide multilayers by pulsed laser deposition and show an enhanced GMR effect in the oxide multilayers compared to the parent $La_{0.7}MnO_{3-\delta}$ magnetic oxide thin films.

Bulk samples of nominal composition $La_{0.7}MnO_{3-\delta}$ and $La_{0.85}MnO_{3-\delta}$ were prepared by heating stoichiometric amounts of La_2O_3 and $MnCO_3$ at 1050°C in air. These oxides in bulk did show ferromagnetic transition at 220 and 240 K, respectively. However, when the films of these materials were made by pulsed laser deposition, they did not show metallic transition. The discrepancy in the properties of the bulk solid and the thin films was attributed to oxygen deficiency. In order to overcome the oxygen out-diffusion in the thin films, 5% Ag_2O was added to the $La_{0.7}MnO_{3-\delta}$ which was used as the laser target. Silver is known to enhance epitaxial growth of $YBa_2Cu_3O_{7-\delta}$ films in addition to bringing in extra oxygen during pulsed laser deposition[8]. Ag ion is not known to substitute for any ion in $La_{1-x}MnO_3$. The films grown with 5% Ag_2O showed metallic conductivity below 220 K. Films made out of $La_{0.85}MnO_{3-\delta}$ pellet (without addition of Ag_2O) showed a semiconducting behaviour which is employed as the nonmagnetic oxide layer in this study.

Thin films of $La_{0.7}MnO_{3-\delta}$ with 5% Ag_2O (Ag-MNO) and $La_{0.85}MnO_{3-\delta}$ (LMO) without Ag_2O were deposited separately on $LaAlO_3$ (1 0 0) substrate by pulsed laser deposition at 720°C in 300 mtorr O_2. Superlattices of $(Ag–LMO/LMO)_{20}$ were also deposited on $LaAlO_3$ (1 0 0) substrate. Ag–LMO and LMO were deposited alternatively using a multitarget holder with 60 and 12 laser pulses, respectively, at a (KrF 248 nm) laser pulse energy density of 2 J/cm^2. From the thicknesses of Ag–LMO and LMO films prepared separately, estimated layer thicknesses in the multilayer are 100 Å and 20 Å, respectively, for Ag–LMO and LMO. X-ray diffraction pattern of the multilayer film given in Figure 1 shows highly textured growth in (1 0 0) direction on $LaAlO_3$ with pseudocubic $a = 3.88$ Å agreeing well with the observed a parameter of the bulk $La_{1-x}MnO_{3-\delta}$. Ag–LMO as well as LMO films also showed a similar oriented growth. Superlattice reflections were not seen in the form of satellite peaks adjacent to (1 0 0) or (2 0 0) reflections of LMO possibly because La^{3+} and oxygen defects are randomly distributed.

Resistivity measured by four point probe method at zero field for a 2500 Å thick epitaxial film of Ag–LMO, LMO and $(Ag–LMO/LMO)_{20}$ multilayer are shown in Figure 2. Sharp change in the slope is observed in the ρ vs T curve, for the Ag–LMO and the multilayer films. The LMO film showed a semiconducting behaviour. Room temperature resistivity of Ag–LMO and LMO

‡For correspondence.

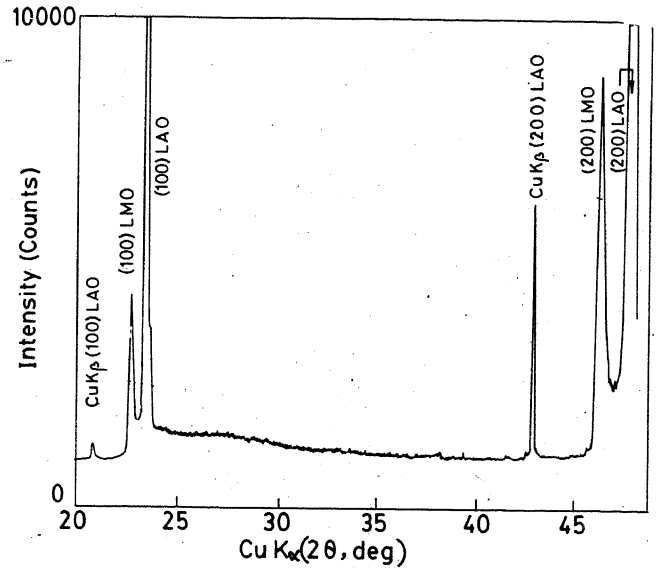


Figure 1. X-ray diffraction pattern of (Ag–$La_{0.7}MnO_{3-\delta}$/$La_{0.85}MnO_{3-\delta}$)$_{20}$ multilayer thin film.

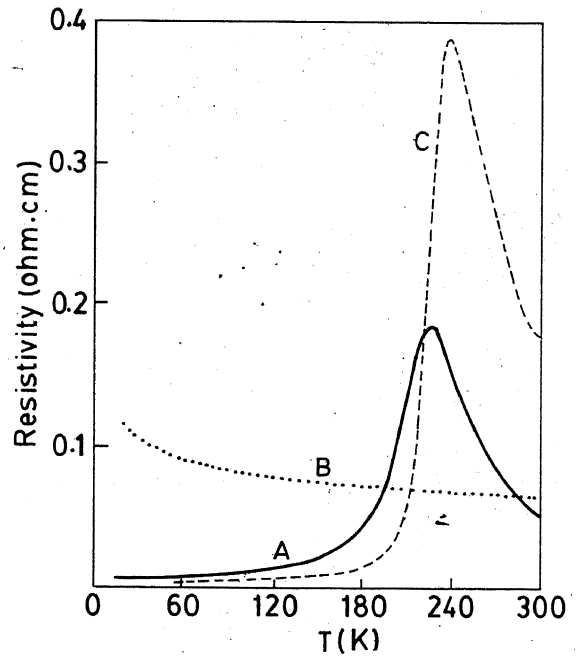


Figure 2. Resistivity vs temperature plots of (A) Ag–$La_{0.7}MnO_{3-\delta}$ (B) $La_{0.85}MnO_{3-\delta}$ and (C) (Ag–$La_{0.7}MnO_{3-\delta}$/$La_{0.85}MnO_{3-\delta}$)$_{20}$ multilayer films on $LaAlO_3$ (100).

films were in the range of 60–70 mΩ cm. Metal to insulator transition temperature, T_c in the multilayer film is about 235 K and that in the single Ag– $La_{0.7}MnO_{3-\delta}$ film is 225 K. The higher resistivity of the multilayer film above 220 K can be accounted for in terms of addition of resistances of the Ag–$La_{0.7}MnO_{3-\delta}$ and $La_{0.85}MnO_{3-\delta}$ films in series. Below the transition temperature the metallic resistivity of Ag–LMO dominates

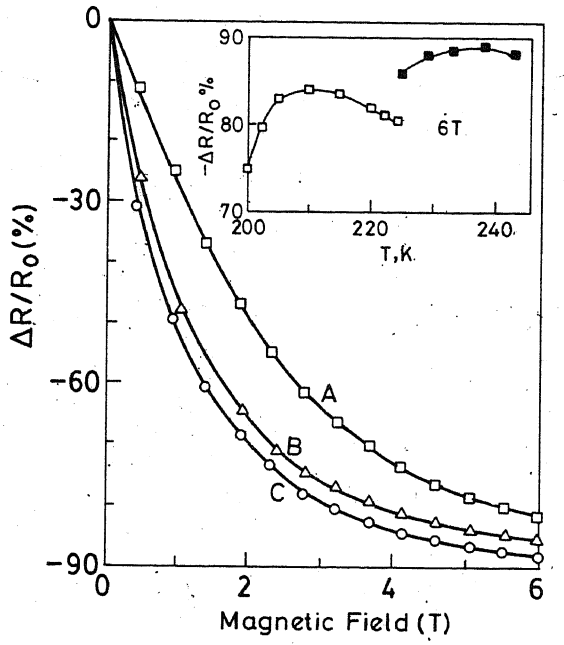


Figure 3. Variation of magnetoresistance $\Delta R/R_0$ as a function of magnetic field at 225 K for (A) Ag–$La_{0.7}MnO_{3-\delta}$ film with H perpendicular to the film; (B) (Ag–$La_{0.7}MnO_{3-\delta}$/$La_{0.85}MnO_{3-\delta}$; (C) Film in B parallel to the field H. Inset shows variation in GMR with temperature at 6 T for Ag–LMO film (□) and multilayer film (■)

and resistivity of the film is essentially a parallel combination of the two oxide layers.

In Figure 3 are shown the percentage $\Delta R/R_0$ vs magnetic field, plots for the parent Ag–LMO (A) and (Ag–LMO/LMO)$_{20}$ (B) multilayer films at 225 K. The measurements were made with H parallel and perpendicular to the film surface. In the case of the parent Ag–LMO film, about 80% GMR is obtained at 6T and there was no difference in the GMR values whether H was parallel or perpendicular to the film surface. A similar isotropic GMR effect is reported in all the doped lanthanum manganate films[3–6]. An enhanced GMR effect is observed in the (Ag–LMO/LMO)$_{20}$ film by about 6–10% compared to the parent Ag–LMO. Further, when the H was parallel to the surface, GMR value was higher by 3% compared to when H was perpendicular (curve C). Anisotropic GMR effect was observed in the case of $(Fe/Cr)_n$ metallic multilayer film as well. In the inset, GMR vs T plots for the parent and multilayer films measured at 6 T parallel to the film surface is shown demonstrating the enhanced GMR effect in the multilayer films.

The pseudocubic cell of doped manganates in general consists of MnO_2 magnetic layers separated by a nonmagnetic La–O or La(M)–O (M = divalent ion) oxide layers. As the temperature is decreased, ferromagnetic order sets in the MnO_2 layer due to Zener double exchange mechanism of electron transfer in the Mn^{3+}–O–Mn^{4+} chains in all the three directions and resistivity

decreases[9,10]. Alternatively, when a magnetic field is applied externally at a given temperature, ferromagnetic order is induced in the MnO_2 layers leading to a decrease in resistance. In the self doped $La_{0.7}MnO_{3-\delta}$, the pronounced GMR effect is attributed to La^{3+} and O^{2-} vacancies which form the scattering centers for the conducting electrons, thereby increasing the resistivity. Applied magnetic field seems to provide sufficient activation energy for the alignment of spins to overcome the additional scattering of conducting electrons.

In the case of multilayer, we propose that there are two factors contributing to the total GMR. The first one is the decrease in the resistances of individual magnetic oxide layers as in the single Ag–LMO film. This value is about 80% at 6T. The second factor is the parallel combination of resistances of the individual magnetic oxide layers separated by the non-magnetic spacer layer. This is about 8–10% thus giving 89% GMR at 6T in the $(Ag–LMO/LMO)_{20}$. The secondary contribution seems to arise from the spin dependent scattering of the conduction electrons at the interface between the ferromagnetic oxide and the nonmagnetic oxide. The order of magnitude of the secondary contribution to the total GMR is about 8–10% which is about the same as that of magnetic/nonmagnetic metal multilayers such as $(Co/Cu)_n$.

In conclusion, we have shown for the first time, a superlattice sequence of magnetic/nonmagnetic oxide multilayer, mimicking the $(Fe/Cr)_n$ superlattices, showing an enhanced GMR compared to the GMR observed in the parent magnetic oxide thin films. We have also shown an anisotropic behaviour in the multilayer oxide films compared to an isotropic behaviour in lanthanum manganate thin films,

1. Searle, C. W. and Wang, S. T., *Can. J. Phys.*, 1970, **48**, 2023–2031.
2. Baibich, M. N., Broto, J. M., Fert, A., Nguyen Van Dau, F., Petroff, F., Etinne, P., Creuzet, G., Fredrick, A. and Chazelas, J., *Phys. Rev. Lett.*, 1988, **61**, 2472–2475.
3. von Helmolt, R., Wecher, J., Holzafel, B., Shultz, L. and Samwer, K., *Phys. Rev. Lett.*, 1993, **71**, 2331–2333.
4. Chahara, K. I., Ohno, T., Kasai, M. and Kozono, Y., *Appl. Phys. Lett.*, 1993, **63**, 1990–1992.
5. Ju, S. L., Kwon, C., Li, Q., Greene, R. L. and Venkatesan T., *Appl. Phys. Lett.*, 1994, **65**, 2108–2110.
6. Sundar Manoharan, S., Vasanthacharya, N. Y., Hegde, M. S., Satyalakshmi, K. M., Prasad, V. and Subramanyam, S. V., *J. Appl. Phys.*, 1994, **76**, 3923–3925.
7. Sundar Monoharan, S., Dhananjay Kumar, Hegde, M. S., Satyalakshmi, K. M., Prasad, V. And Subramanyam, S. V., *J. Solid State Chem.* (in press).
8. Dhananjay Kumar, Satyalakshmi, K. M., Sundar Monoharan, S. and Hegde, M. S., *Bull. Mater. Sci.*, 1994, **67**, 625–632.
9. Goodenough, J. B., *Progr. Solid State Chem.*, 1971, **5**, 149–350.
10. Mahesh, R., Mahendiran, R., Raichaudhury, A. K. and Rao, C. N. R., *J. Solid State Chem.*, 1995, **114**, 297–299.

ACKNOWLEDGEMENTS. We thank the Department of Science and Technology, Government of India, for financial assistance. KMS thanks CSIR for a senior research fellowship.

Received 31 March 1995; revised accepted 10 July 1995

Physiobiochemical aspects of development of hardseededness in *Albizia lebbek*

R. C. Uniyal and A. R. Nautiyal

High Altitude Plant Physiology Research Centre,
HNB Garhwal University, Srinagar Garhwal 246 174, India

Seeds of *Albizia lebbek* exhibit dormancy due to hard seeded coat. Seed development was studied in this species to determine whether the induction of hardseededness precedes or succeeds the attainment of physiological maturity and also whether the hardseededness is an inherent character of the seed or can be escaped by harvesting seeds before its induction. The seeds attained physiological maturity at approximately 230 DAA accompanied by acquisition of germinability. This phase was characterized by rapid accumulation of starch, soluble carbohydrates, proteins and chlorophyll. The onset of hardseededness followed, proceeding gradually until 270 DAA and then increased rapidly. While the soluble carbohydrates and chlorophyll contents declined during this period, the starch and proteins showed an increase. Hardseededness in these seeds seems to be their inherent character as it not only occurred in seeds attached to the parent plant but also in seeds in storage, although considerably delayed in the latter case.

DELAYED germination of seeds due to hard seed coats is a common characteristic of many leguminous seeds. *Albizia lebbek* (L.) Willd. is one such multipurpose tree species in which hardseededness poses a practical problem for germination of its seeds. The occurrence of hardseededness in some seeds indicates that the prevailing environment during seed development plays a significant role in developing a hard seed coat[1–3]. However, information is lacking on the germinability and performance of seeds harvested prior to induction of hardseededness. Thus, the present investigation was undertaken to determine whether seeds attain germinability before the induction of hardseededness, and if so, whether hardseededness could be escaped by harvesting seeds before its induction.

Table 1. The physiobiochemical changes and germination in the seeds of *A. lebbek* during development (Mean ± SE)

Stage (DAA)	Pod fr wt (g)	Seed wt. (mg) Fresh	Seed wt. (mg) Dry	Seed MC	Chl a+b mg/g fresh weight	Sol.sug mg/g fresh weight	Starch mg/g fresh weight	Sol. prot. mg/g fresh weight	Percentage germination after storage (months)* 0	1	2	3	4	5
M_1 (150)	7.28	39.39 ±1.38	7.58 ±0.27	80.92 ±0.17	0.163 ±0.001	32.52 ±0.08	10.33 ±0.005	53.99 ±1.53						
M_2 (170)	7.25	154.81 ±4.45	32.87 ±0.94	78.79 ±0.18	0.154 ±0.002	39.80 ±1.42	23.46 ±0.21	76.05 ±1.43						
M_3 (190)	8.37	231.99 ±5.05	68.72 ±1.78	70.39 ±0.37	0.143 ±0.01	33.42 ±0.68	115.81 ±3.43	127.66 ±1.17	80.00					
M_4 (210)	6.42	272.86 ±12.74	93.13 ±4.76	65.91 ±0.31	0.283 ±0.004	39.28 ±0.38	195.71 ±13.01	177.77 ±1.18	100.00					
M_5 (230)	3.85	177.01 ±5.12	89.50 ±2.59	49.37 ±0.81	0.032 ±0.002	75.39 ±9.33	191.93 ±13.62	241.67 ±6.35	94.67	85.00	81.67	88.33	75.00	10.00
M_6 (250)	3.89	134.38 ±3.27	80.76 ±2.44	39.75 ±1.78	0.081 ±0.002	55.17 ±5.19	125.87 ±16.40	255.24 ±1.15	87.00	77.33	ND	–	–	–
M_7 (270)	3.57	112.59 ±1.87	82.91 ±1.91	26.22 ±0.67	0.027 ±0.0003	56.81 ±3.09	163.94 ±4.80	288.81 ±9.49	85.33	81.33	65.33	60.00	60.00	21.66
M_8 (290)	3.67	111.09 ±1.29	85.79 ±1.53	22.81 ±0.68	0.029 ±0.001	50.63 ±1.68	174.60 ±5.82	355.55 ±1.16	65.00	55.00	35.00	13.33	8.33	0.00
M_9 (310)	2.95	117.18 ±2.84	99.96 ±3.27	14.80 ±0.98	0.030 ±0.0005	58.96 ±3.79	210.01 ±12.50	339.01 ±24.25	0.00	0.00	11.67	10.00	10.00	10.00

DAA – Days after anthesis.
ND – Not determined due to accidental loss of seeds.
* – These data represent percentage germination without any treatment. After hot water treatment the germination ranged from 90 to 100%.

The study was undertaken in two subsequent years 1991–1992 and 1992–1993. Careful monitoring of pod and seed development was done beginning at the full-bloom stage and commencing 150 days after anthesis (DAA) when seeds were still minute in size but yet easily separable from the pod. The observations on the development of pod and seeds of *A. lebbek* revealed that development of seeds was delayed until pods had attained almost full growth, after which seeds began to develop rapidly. It required 150 days from anthesis for pods to reach this stage. Thus, pods of *A. lebbek* were collected from a natural stand at Srinagar Garhwal (550 m altitude) at 20-day intervals beginning 150 DAA and continuing until 310 DAA when pods reached a dry stage and started shedding. Stages were identified as M_1, M_2, M_3 310 DAA. Following each harvest, ten randomly chosen pods were analysed for their weight, size and colour. Fifty pods were dissected to liberate their seeds, which were mixed together and analysed for their weight, moisture content, biochemical analyses, germination and storability.

The moisture content of ten randomly chosen seeds was determined by drying them at 80°C for 48 h in an electric oven, after first determining their fresh weight, and calculating by the formulae given by Evans[4]. For germination, seeds were kept in three sets, each set containing 50 seeds, on the top of Whatman no. 1 filter paper in Petri dishes of 9 cm diameter and moistened daily with distilled water. Petri dishes were placed at room temperature, which varied between 17 and 24°C during the course of study. Daily observations were made for monitoring the germination (radical emergence) up to 30 days after sowing. To test the germinability of hard seeds, the seeds were soaked in hot water before placing them for germination. To study the changes in composition during maturation, the contents of chlorophyll, soluble sugars, starch and soluble proteins in freshly harvested seeds were determined. For chlorophyll determination Holm's[5] method, for carbohydrates the Anthrone method[6] and for soluble protein estimation the method given by Lowry *et al.*[7] were used.

As mentioned above, development of seed was delayed until pods attained almost full growth, which took 150 days after anthesis. Pod growth continued for an additional 40 days (M_3 stage) and subsequently ceased, after which the fresh weight of pods started declining (Table 1). In contrast, the fresh weight and dry weight of seeds increased progressively for 60 days until 210 DAA (M_4 stage), and subsequently a significant decline in seed fresh weight was recorded, which reached less than 50% of its maximum fresh weight at the time of seed shedding. However, after reaching a maximum on 210 DAA, the dry weight decreased slightly, which was not statistically significant except in the last stage. The moisture content of the seed started declining 170 DAA, first slowly and then very rapidly. Moisture content declined to 50% at 230 DAA, finally reaching 14.8% on 310 DAA. Maximum pod length and breadth were attained between 170 and 190 DAA; thereafter, while the length remained static, the breadth displayed gradual attenuation. The pod, however, exhibited a series of colour changes from green up to 230 DAA,

Table 2. Coefficient of correlation between various parameters studied during seed development in *A. lebbek* ($n = 10$ in each case)

	Days after anthesis	Seed dry weight
Starch	0.792****	0.969*****
Sugar	0.650**	0.624*
Protein	0.985*****	0.822****

Levels of significance.
$p = 0.1$*, $p = 0.05$**, $p = 0.01$****, $p = 0.001$*****.

then greenish yellow between 230 and 270 DAA, and finally a dark yellow.

The chlorophyll content was 0.283 mg g^{-1} until 210 DAA. Thereafter, it declined to 0.030 mg g^{-1} fr. wt at the time of pod shedding. Contrary to this, the soluble sugar content, starch and soluble proteins increased during seed development. Statistical analysis of the data shows a positive and highly significant correlation in biochemical constituents and seed development (Table 2). While a decrease was observed for carbohydrates at the maximum seed weight stage, soluble protein content continued to increase until the second last stage. During this period of seed development, a remarkable increase in all biochemical constituents was recorded during the early phase, i.e. between 150 and 230 DAA. It was during this period that seed attained maximum fresh and dry weight. Starch, soluble carbohydrates and soluble proteins all had a positive correlation with seed dry weight, which was highly significant in the case of starch ($p = 0.001$) and proteins ($p = 0.01$), indicating that these are the major contributors for the seed dry weight. The increasing trend of soluble sugars, starch and soluble proteins during seed maturation showed the relationship between the synthesis of nitrogenous compounds, i.e., proteins, during maturation (important in the formation of organelles, cytosol and membranes) and the metabolites that participate in vital processes in the seeds, as reported by Chamma *et al.*[8].

The developing seeds attained germinability around 190 DAA as 80% germination was obtained in seeds at this stage without exhibiting any hardseededness. By 210 DAA all seeds had acquired the ability to germinate as 100% germination was recorded at this stage (Table 1). This was the stage when the mean dry weight of the seed was also maximum. Physiological maturity in angiosperms has generally been designated as completion of organic accumulation within the seed as reflected by no further increase in its dry weight[9, 10]. Thus, in the present study it seems that seeds of *A. lebbek* attained physiological maturity around 210 DAA, because seed dry weight was maximum and germination was 100%, but the ability to germinate even prior to this stage indicates that acquisition of germinability preceded attainment of physiological maturity in this case, as increase in seed dry weight between M_3 and M_4 stages was statistically significant. However, this is not an exception, as in some other legumes also the seeds developed the ability to germinate even prior to the attainment of maximum dry weight[11]. Subsequent to 210 DAA, a gradual decline in percentage germination was recorded which was remarkable during the final stage between 290 and 310 DAA, when a 65% decline was recorded in the germinability of seeds and thus the seeds seemed to have attained complete hardseededness at this stage as no germination was recorded without scarification. Thus, there is a clear cut indication that hardseededness in *A. lebbek* is induced after attainment of physiological maturity, characterized by maximum seed dry weight, the M_4 stage. In spite of the onset of hardseededness after the M_4 stage, a high percentage of seeds still exhibited germination ranging from 95% at M_5 stage to 85% at M_7 stage without any scarification. This may be due to the slow progress in development of hardseededness during this period; therefore, seeds can be harvested at these stages.

Seeds harvested at M_4 stage could not be stored viably but those from subsequent stages not only stored well but also exhibited an interesting pattern of germination during storage. Seeds stored at room temperature for five months and tested for their germinability at monthly intervals showed a gradual decline in their percentage of germination with their age in storage, not because of the loss of viability but due to development of hardseededness, which was evidenced by their germination after hot water treatment. This decline in germination was faster in the seeds harvested at advanced stages (e.g. M_8 or M_7 stage) than at an early stage (M_5 stage), as evident by the data in Table 1. After four months in storage, while percentage germination declined only by 20% as compared to freshly harvested seeds at M_5 stage, in case of M_7 and M_8 stages the decline was 30 and 87%. By six months all seeds in storage had developed hardseededness and none germinated without scarification. Thus, hardseededness seems to be an inherent character of these seeds and is not related to its attachment with the parent plant as the process of development of hardseededness continued even after the seed had been detached from the plant and kept in storage. Adams and Rinne[12] have also reported for *Glycine max* that the maturation is an inherent function of the seed, the completion of which does not depend on the association of the seed with the parent plant. Interestingly, here the development of hardseededness was faster when the seeds were attached to plant than in storage.

The evaluation of parameters that allow the visual identification of physiological maturity or some information that contributes significantly to decisions concerning harvest of the seeds is another very important aspect of seed development studies. In many hardwoods, cone/fruit colour was investigated as a workable indicator for maturation, e.g. in five species of *Quercus*[13, 14]. A series of investigations[8] carried out in many agricultural crops, e.g. corn, sorghum, oat, soybean, etc., revealed that the formation of the black layer in corn and sor-

ghum and the changing colour of the fruits and seeds in other species allowed an accurate identification of the physiological maturity of seeds. According to results obtained here, on the basis of pod colour changes in *A. lebbek* it can be concluded that the colour of the pods was closely related to the maximum physiological quality of seeds as they started changing colour from green to yellow around 230 DAA, when seeds had attained physiological maturity. Thus, the time when green pods are not found on the plant any longer can be used to characterize the physiological maturity of seeds, and pods can be collected at this stage. Thus, the overall results obtained here indicate that development of hard-seededness in seeds of *A. lebbek* succeeds the attainment of physiological maturity, and it is an inherent character of the seed which can, however, be delayed to some extent by harvesting the seeds immediately after maturation. The colour of the pods served as a workable indicator of maturation.

1. Donelly, E. D., *Crop Sci.*, 1970, **10**, 661–662.
2. Kumari, P. and Dahiya, B. S., 1st National Workshop on Seed Technology at H. A. U., Hisar, Haryana, 1981.
3. Tomer. R. P. S. and Kumari, P., *Seed Sci. Technol.*, 1991, **19**, 51–56.
4. Evans, G. C., *The Quantitative Analysis of Plant Growth*, Blackwell, Oxford, 1972.
5. Holm, G., *Acta Agric. Scan.*, 1954, **4**, 457.
6. McCready, R. M., Guggloz, J., Silviera, V. and Owens, H. S., *Anal. Chem.*, 1950, **22**, 1156–1158.
7. Lowry, O. H., Rosebrough, N. J., Farr, A. L. and Randall, R. J., *J. Biol. Chem.*, 1951, **192**, 265–275.
8. Chamma, H. M. C. P., Marcos-Filho, J. and Crocomo, O. J., *Seed Sci. Technol.*, 1990, **18**, 371–382.
9. Rajana, B. and Andrews, C. H., *Proc. Assoc. Offi. Seed Anal.* 1970, **60**, 188–196.
10. Silva, C. M., Vieira, C. and Sedyiama, C. S., *Semente (Brasilia)*, 1975, **1**, 12–20.
11. Ellis, R. H., Hong, T. D. and Roberts, E. H., *Ann. Bot.*, 1987, **59**, 23–29.
12. Adams, C. A. and Rinne, R. W., *J. Exp. Bot.*, 1981, **32**, 615–620.
13. Bonner, F. T., *Forest Sci.*, 1974, **20**, 238–242.
14. Bonner, F. T., *Forest Sci.*, 1976, **22**, 149–154.

ACKNOWLEDGEMENTS. The financial assistance from MNES to ARN and CSIR to RCU is gratefully acknowledged.

Received 1 May 1995; revised accepted 6 July 1995

Fractal description of seismicity of India and inferences regarding earthquake hazard

K. N. Khattri

Wadia Institute of Himalayan Geology, Dehradun 248 001, India

Earthquakes have the quality of fractal structure in their spatial disposition, time sequencing, and magnitude distribution. We analyse the seismicity of India in terms of its spatial fractal structure in various seismic source zones. It is found that the fractal dimensions range between 0.894 and 1.574, indicating that at most the earthquake associated fractures are approaching a two-dimensional space. The low fractal dimension is suggestive of the presence of asperities and barriers of significant sizes in the respective source zones. Using a fractal model the average earthquake slip in primary faults is estimated to be about 3.5 cm/yr in the Himalaya. This gives an average return period of great earthquake of about 285 years. Considering that no great earthquake has occurred in the seismic gaps in the Himalaya for at least 200 years, these gaps are ripe for a future great earthquake to occur within the next hundred years.

To see a world in a grain of sand,
And a Heaven in a wild flower,
Hold Infinity in the palm of your hand,
And Eternity in an hour.

– William Blake

THE earthquakes represent the outcome of complex geomechanical processes having dire consequences for the society. The earth is in an overall steady state of strain, which for most part is caused by plate motions. This is released from time to time in the form of catastrophic motion over faults in the crust, resulting in earthquakes. The structure of the crust is highly inhomogeneous. On a microscale there are crystal defects, cleavages, dislocations, grain boundaries, etc., while on the other end of the scale, i.e. macroscale, by virtue of repeated fracturing and the different shapes of fractures a high degree of heterogeneity of the crust has evolved. The strain field accordingly is also quite heterogeneous. On account of such complexity of the medium, precise prediction of the occurrence of earthquakes in terms of their location, magnitude and time of occurrence is not feasible. Therefore, the earthquakes are describable by statistical means.

Mandelbrot[1] observed that many natural phenomena possess self-similarity at many different scales. He used the term 'fractal' to describe such phenomena. For example, an assemblage of objects with differing sizes and irregular shapes is a fractal set if the number of objects in it with a specified size has a power law dependence on size. The scaling parameter is called the fractal dimension D. The geometry of the fracture surface of rocks is a fractal[1, 2]. The natural rock fracture process is a fractal. Earthquakes are associated with fractals by virtue of the accompanying fracturing of rocks. A fractal distribution is scale-invariant. In such a set the complexity of the part is as great as that of the whole. A fractal function such as a curve or a surface is not dif-

ferentiable and, thus, no local rule can be deduced to generate the function. However, inferences can be drawn about the whole from observations of the phenomenon on any scale; this is the basis of a holistic approach for studying fractal systems[3]. In comparison, in an ordinary system, howsoever complex it may be as a whole, the basic elements are believed to be simple, which is the basis of a reductionist approach.

Earthquakes possess fractal structure with respect to time, space and magnitude. The Gutenberg–Richter relation for frequency vs magnitude is a power law. Similarly, the aftershock decay follows another power law involving time[4]. The two-point spatial correlation function for earthquake epicentres also displays a power law structure[5–7].

Aki[8] proposed that fractal dimension is correlated with the b value of the Gutenberg–Richter relation. The fractal dimension could be interpreted in terms of the nature of crustal deformations accompanying slips on faults, which may be governed by the sizes and spatial distributions of asperities or barriers in the crust[9]. King[10] and Turcotte[11] have used Aki's model to interpret the crustal deformation due to slips on fractal fault systems. However, Hirata[12] has found that for Japan this idea is not satisfactory.

In this paper we use Aki's[8] model to determine the fractal dimension of seismicity in various source zones in the Indian region and make comparative inferences regarding fault segmentation, asperity–barrier relationships and consequences for seismic hazard estimation. We further use Turcotte's[11] formalism to estimate the fraction of slip taking place on the main fault, the rest being accommodated on a smaller fault system. Thus, we estimate the strain rate and make prognostications about the average return periods of largest earthquakes.

The Gutenberg–Richter frequency magnitude relation is given by

$$\log N\,(m > M) = a - bM, \tag{1}$$

where N is the cumulative number of earthquakes in a time window in an area having magnitudes m greater than M, a and b are parameters of correlation. This relation allows us to estimate the average return period of a future earthquake of a given magnitude, and is, therefore, the basis of preparing probabilistic seismic hazard maps[13].

The Gutenberg–Richter relation represents a power law distribution of magnitudes. Since magnitudes are related to the fault size involved in the earthquake generation, a power law relation for the faults causing earthquakes may be inferred[3].

Another power law relation has been obtained between the magnitude M (surface wave) and the moment M_0 of an earthquake[14]:

$$\log M_0 = d - cM, \tag{2}$$

where c and d are parameters of correlation.

Let both L and l represent linear dimensions. Assuming that the distribution of the earthquake sizes is self-similar and, therefore, the seismic moment M_0 is proportional to L^3 (volume), Aki[8] has shown that a relation, between the cumulative number of faults $N(l > L)$ and L, of the following type may be obtained:

$$\log N(l > L) = e - (3b/c) \log L, \tag{3}$$

where e is a constant.

The power law relation discussed above characterizes the earthquake process. It is appropriate to examine it with the analytical framework for fractals. The parameter characterizing a fractal set is its fractal dimension D.

Let N_n be the number of fragments with characteristic linear dimension r_n in a fractal set; then a power law of the type given below exists for the set[15]:

$$N_n = f r_n^{-D}, \tag{4}$$

where f is a constant of proportionality and D is the fractal dimension. D lies between D_T (topological) and D_E (Euclidean) and may be a fractional number.

Comparing (3) and (4) the fractal dimension for the earthquake process is $3b/c$. An average value of c is estimated to be 1.5 (ref. 13). Thus, on an average, $D = 2b$.

The value of b in various seismogenic zones of the earth has been found to range between about 0.5 and 1.5. Thus, the corresponding fractal dimension can range between 1 and 3. A value close to 3 implies that earthquake fractures are filling up a volume of the crust, a value close to 2 suggests that it is a plane that is being filled up and a value close to 1 means that line sources are predominant[8].

Slip over faults distributed over a region form a major mode of crustal deformation. At the major plate boundaries this slip is driven by relative velocity between the plates. Of the total slip a fraction occurs over the primary fault system causing the largest (and also smaller) earthquakes and the rest on the secondary fault system[11]. An estimate of the proportion of the slip occurring over the primary faults can permit a better estimation of the return periods of the largest earthquakes.

Representing symmetric slip at triple junctions by rhombic plates that together form a regular hexagon, King[10] showed that for fractal distribution of such faults the corresponding earthquakes follow the Gutenberg–Richter relation. Using this concept of fractal matrix of faults Turcotte[11] represents the slip over an extended fault by superposition of slips over the primary fault and that on a fractal distribution of self-similar system of smaller faults. For such a scheme, the following relation for the ratio of the slip over the primary fault to the total slip over the fault system (i.e. due to the relative motion of the plates with respect to each other) is given:

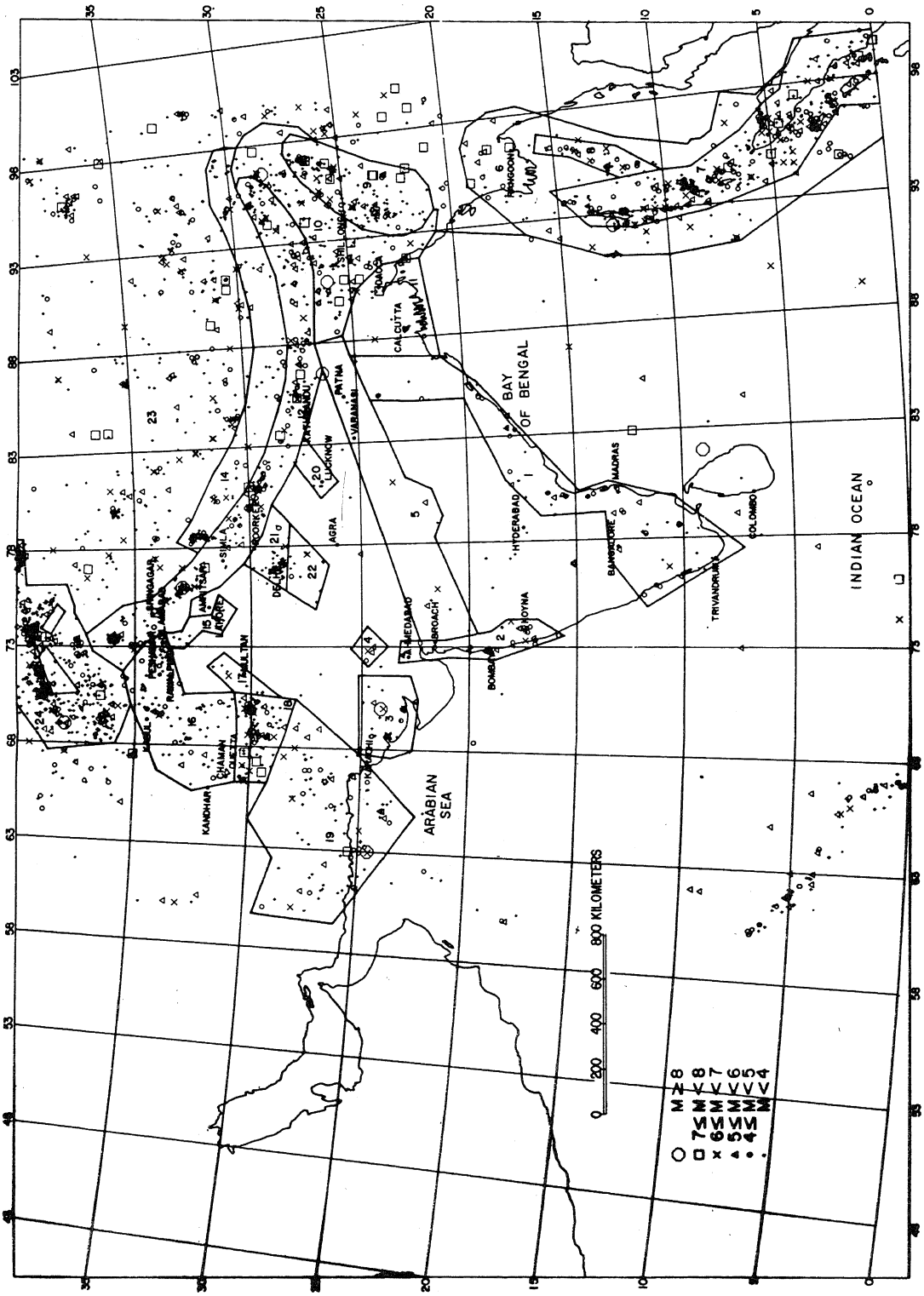


Figure 1. The seismic source zones of Indian region (after Khattri *et al.*[13]).

$$s_p/s = 1 - 2^{-(3-D)}, \quad (5)$$

where s_p is the slip on the primary fault, s the total slip and D the fractal dimension. For $D = 3$ all the slip occurs on smaller faults, whereas on the other end, for the limiting case of $D = 0$, nearly 88% of the slip occurs on the primary faults. The slip that occurs on a fractal system of subfaults contributes to gradual processes like fold formation and other similar crustal deformations.

The seismicity of the Indian region has been examined in detail by Khattri *et al.*[13]. They identified earthquake source regions on the basis of quasi-homogeneity of seismotectonic processes in the respective source regions (Figure 1). The seismicity catalogue was examined and compensated for incompleteness, and the b values were estimated. In some source zones, those in Peninsular India, the seismicity data were not adequate for an independent and reliable estimate of the b values to be made. For such source zones respective data sets were combined and an estimate of the b values was made. The estimates of the b values and the corresponding D values are shown in Table 1.

We note that the fractal dimensions for various source zones range between 0.894 for the Shillong region and 1.574 for the Peninsular region.

The fractal dimensions estimated above indicate that the fractures in the various source zones are filling up either linear regions (low fractal dimensions of ~1) or nearly two-dimensional surfaces (fractal dimension of ~1.5). None of the source zones has a volume distribution of sources.

The largest b value of 0.787 is estimated for the source zones of Peninsular India, for which, as mentioned above, the earthquake catalogues were combined. The compositing of the catalogues of earthquakes for various source zones, possibly having different geodynamic processes, may have a biasing effect towards higher estimates for the b value, as we find in the present analysis. This could be a reason why we get a b value estimate of nearly 1 for the global catalogue of earthquakes. In this light the estimates of D for source zones in the Peninsular region are not reliable.

The low b value and the corresponding low fractal dimensionality obtained for the Shillong region (zone 10) is noteworthy and indicates that there is a tendency for few faults with large attendant earthquakes on them. This may be interpreted to be indicative of a relative lack of fracturing and attendant weaker zones in the region. In other words, the region is characterized by the presence of large asperities or barriers which are relatively unfractured and strong. Similarly, the low fractal dimensionalities of source zone 6 in the Andaman region and 18 in the Sulaiman–Kirthar region imply the presence of significant asperities or barriers.

The Himalaya (source region 12) and the source zone 14 to the north have fractal dimensions of 1.514 and 1.226, respectively. This suggests that the fracture distributions are tending to fill a two-dimensional (2-D) space.

Table 1. The b and D values for the seismic source zones

Source zone	b value	D value	Region
1–5	0.787	1.574	Peninsula
6	0.533	1.066	Andaman
7	0.678	1.356	Andaman
8, 9	0.580	1.160	Andaman, Arakan
10	0.447	0.894	Shillong
11	0.787	1.574	Bengal Basin
12	0.757	1.514	Himalaya
14	0.613	1.226	Trans Himalaya
15	0.787	1.574	Himalaya foredeep
16	0.581	1.162	Sulaiman
17	0.787	1.574	Sulaiman foredeep
18	0.485	0.970	Sulaiman–Kirthar
19	0.581	1.161	Kirthar–Makran
20, 21	0.787	1.574	Himalaya foredeep
22	0.787	1.574	Peninsula
23	0.543	1.086	Tibet
24	0.561	1.122	Hindukush

The other source zones are characterized by fractal fracture distributions ranging in between the two end types discussed above.

Following Hirata[12], we calculated the 2-D fractal dimension D_2 using the correlation integral $C(r)$ defined with respect to the spatial distribution of epicentres. It is defined by

$$C(r) = [2/\{N(N-1)\}]N(R < r), \quad (6)$$

where $N(R < r)$ is the number of pairs of epicentres with a mutual distance R smaller than r. N is taken to be 865 events. The value obtained is 0.275. Now D (in Aki's definition) is the volume dimension (3-D). Thus,

$$D = D_2 + 1,$$

from which we obtain a value of 1.275 for D. The value obtained from the b value is 1.514. The two estimates are reasonably close. As noted above, it was not so in the case of the Japanese earthquakes[12]. Thus, we consider that the present analysis is valid.

We next consider the estimation of the slip on the primary fault in the Himalayan collision zone. The convergence rate between India and Asia is estimated to be 5.5 cm/yr. Of this a fraction is released on the primary fault and the rest on the fractal distribution of subfaults. The primary fault is considered to be the basement thrust[16], on which the largest earthquakes occur. This is a north-dipping surface constituting a megathrust extending along the Himalaya for about 2400 km and having an average width of approximately 80 km. It is desirable to estimate the fraction of the total convergence rate released over this fault, as that would enable one to estimate the average return period of great damaging earthquakes. In the present analysis the source zones 12 and 14 are considered to be the main players in

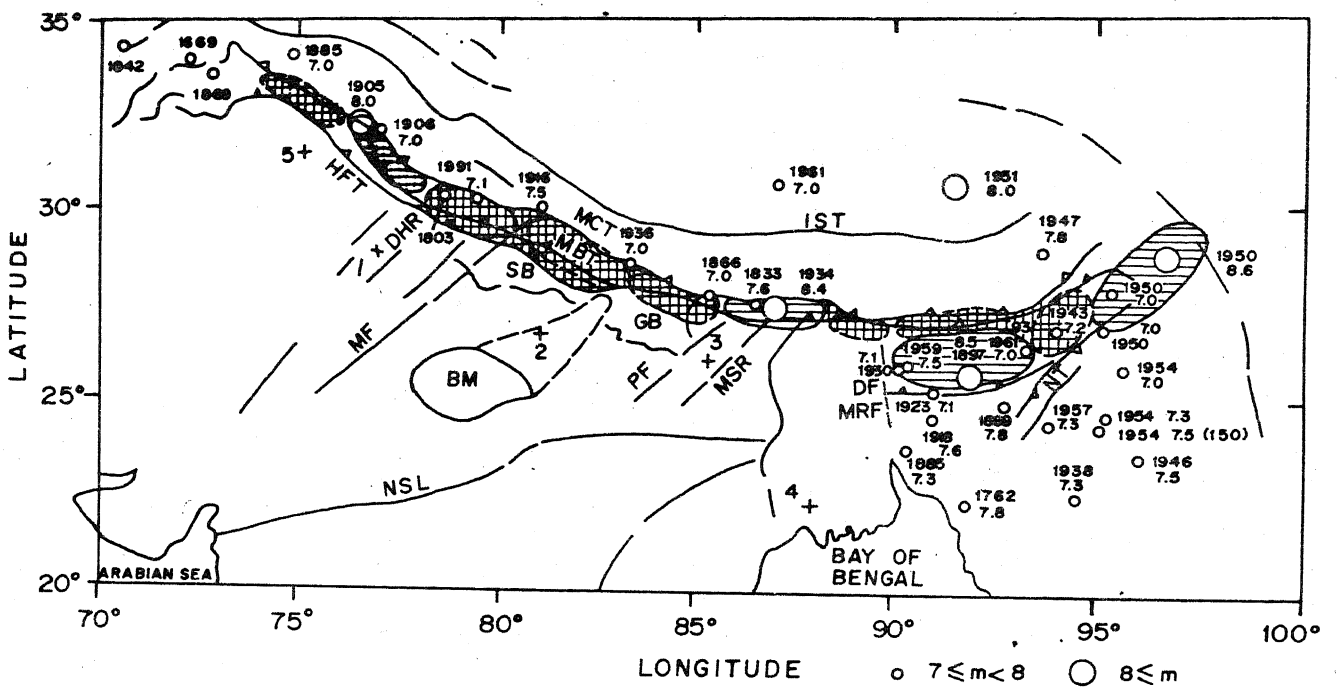


Figure 2. The larger earthquakes of Himalaya. The rupture zones of the great earthquakes are shown by hatured areas. The outlines of the sections of the seismic gaps that may rupture in a single earthquake are marked by dashed lines. It is possible that the central seismic gap may rupture in a single great earthquake[20]. 1: Delhi; 2: Lucknow; 3: Patna; 4: Calcutta; 5: Lahore; BB: Bundelkhand massif; DF: Dauki fault; DHR: Delhi–Hardwar ridge; GB: Gandak basin; HFT: Himalaya frontal thrust; MF: Moradabad fault; MRF: Madhupur fault; MBT: main boundary thrust; MCT: main central thrust; MSF: Monghyr–Saharsa ridge; NSL: Narmada Son lineament; NT: Naga thrust; PF: Patna fault; SB: Sarda basin (after Khattri[22]).

releasing the strain. Further, the source zone 14 comprises of only the fractal subfaults, the primary fault being confined to the source zone 12. Since the fractal dimensions in the two source zones are quite close, we take the value for the same from source zone 12. Using this value in equation (5) we obtain a value for slip rate s_p on the basement thrust to be approximately 3.5 cm/yr. The remaining slip is released on the smaller fractally distributed subfaults. Thus, on an average, for a magnitude 8+ earthquake involving 10 m of slip a time interval of 285 years will be required. The average slip rate for underthrusting is estimated to be of the order of 1.0–2.5 cm/yr (ref. 17) from moments of large earthquakes and geological evidence, 1.5–2.2 cm/yr (ref. 18) from GPS measurements in Nepal section of the Himalaya, and 7 ± 4 cm/yr (ref. 19) from geodetic measurements using levelling. All these estimates are of the same order of magnitude, with the estimate for the slip rate obtained here being somewhat higher. If part, say, one-third, of the slip on primary faults is taken up in source zone 14, then the estimate obtained here will be almost totally congruent with the other estimates. We note that while the geological estimate is an average over several million years, the estimate from GPS data is instantaneous. One of the two estimates derived from earthquake data is based on a period of most recent, about a couple of hundred years, data and magnitudes ranging from smaller to largest earthquakes, while the other one is derived on the basis of great earthquakes only. The estimate from levelling data refer to a period of a few decades. The consistency of estimates derived by such a diversity of databases as well as concepts serves to instill a measure of confidence in the estimated value.

The Himalayan plate boundary has experienced four great earthquakes ($M > 8$) in 1897, 1905, 1934 and 1950, which ruptured and relieved accumulated strains over about 200–300 km long sections (Figure 2). These earthquakes have left sections of the plate boundary where the accumulated strains have not been relieved for at least 200 years; these sections form the seismic gaps[16–26]. The large earthquakes in the Himalaya are recorded from 1803 onwards and a total of 15 with magnitudes (estimated from damage reports for earthquakes in the previous century) between 7.0 and 7.6 have occurred. Of these, two are aftershocks of the great 1950 Assam earthquake and one of the great 1905 Kangra earthquake. One of magnitude 7.6 occurred in 1866 in the rupture zone of the subsequent great 1934 Bihar–Nepal earthquake. Four events, in 1923, 1930, 1959 and 1960 have occurred in the rupture zone of the great Assam earthquake of 1897 and may be considered to be its aftershocks. One event has occurred in 1943 in the Assam gap.

An 800 km long central seismic gap exists between the rupture zones of the Kangra and the Bihar earthquakes. It is possibly divided by transverse geological structures into three sectors. The first lies between the Delhi–Hardwar ridge (eastern margin of the 1905 Kangra rupture zone) and the Moradabad fault, the next is between the Moradabad fault and the Bundelkhand massif promontory, and the third one is between the promontory and the Patna fault. The first sector has had earthquakes in 1803 and 1991. One event occurred in line with the Moradabad fault, another event occurred in 1866 in a region aligned with the Patna fault. The average slip in a $M \sim 7$ earthquake is of the order of 1–2 m. This has been released over faults having relatively small areas, in various parts of the central sector, through such earthquakes since 1803. The cumulative fault area (overall $M \sim 7$ earthquakes) is estimated to be merely about 4500 km^2. On the other hand, the seismic gap has accumulated a slip of at least 5–7 m over an area of 48,000 km^2 (length of about 800 km). This slip accumulation would be much more in case the last great earthquake in the gap occurred 300 or more years ago. Thus, this and similarly the Assam gap have very high potential for giving rise to great earthquakes in the not too distant a future (within about 100 years); depending on how far back in time the last great earthquake ruptured these seismic gaps, the next one may be advanced in time accordingly. In this connection it may be noted that the earliest report of a damaging earthquake in northwest Himalaya appeared in 1669 and in Nepal Himalaya in 1833 ($M = 7.6$). Thus, we may reasonably conclude that an earlier great earthquake in the Himalaya has occurred prior to the 18th century. This already gives at least 300 years of strain accumulation.

The seismic zone of the Himalaya has a fractal dimension lying between 1 and 2, and may be taken to behave approximately as a planer zone. On the other hand, the Shillong massif region has a fractal dimension of less than 1. Thus, this region is relatively dominated by the presence of asperities and/or barriers.

The results of the present study supply confirmatory evidence supporting the conclusions arrived at by earlier studies[16–26] that the seismic gaps in the Himalaya have considerably increased potential for the great earthquakes to occur within the next 100 years or so.

Therefore, there is an urgent need to recognize the above findings regarding the lurking earthquake disasters. The public and government opinions need to be mobilized for a well-planned effort for intensive earthquake studies in the Himalaya, which at present are at a very rudimentary plane. Also there is an urgent need for mounting on a war footing a systematic earthquake hazard mitigation plan to ameliorate the disastrous consequences from future earthquakes.

1. Mandelbrot, B. B., *The Fractal Geometry of Nature*, Freeman, San Francisco, 1982.
2. Brown, S. R. and Scholtz, S. H., *J. Geophys. Res.*, 1985, **90**, 12575–12582.
3. Ito, K., *Pageoph.*, 1992, **138**, 533–548.
4. Omori, F., *J. Coll. Sci. Imp. Univ. Tokyo*, 1894, 7, 111–200.
5. Kagan, Y. Y. and Knopoff, L., *Geophys. J. R. Astr. Soc.*, 1980, **62**, 697–717.
6. Sadvskiy, M. A., Golubeva, T. V., Pisarenko, V. F. and Shniraman, M. G., *Izy. Acad. Sci. USSR. Earth Phys.* (Engl. Transl.), 1984, **20**, 87–96.
7. Ogata, Y., *J. Am. Stat. Assoc.*, 1988, **83**, 9–27.
8. Aki, K., in *Earthquake Prediction: An International Review*, M. Ewing Series (eds Simpson, D. W. and Richards, P. G.), Am. Geophys. Union, Washington, DC, 1981, vol. 4, pp. 566–574.
9. Lay, T. and Kanamori, H., in *Earthquake Prediction: An International Review*, M. Ewing Series (eds Simpson, D. W. and Richards, P. G.), Am. Geophys. Union, Washington, DC, 1981, vol. 4, pp. 579–592.
10. King, G., *Pageoph.*, 1983, **121**, 761–815.
11. Turcotte, D. L., *Pageoph.*, 1986, **132**, 361–369.
12. Hirata, T., *J. Geophys. Res.*, 1989, **94**, 7507–7514.
13. Khattri, K. N., Rogers, A. M., Perkins, D. M. and Algermissen, S. T., *Tectonophysics*, 1984, **108**, 93–134.
14. Kanamori, H. and Anderson, D. L., *Bull. Seismol. Soc. Am.*, 1975, **65**, 1073–1095.
15. Turcotte, D. L., *Fractals and Chaos in Geology and Geophysics*, Cambridge University Press, New York, 1992, 221pp.
16. Seeber, L. and Armbruster, J. G., *Geodynamic Series 3*, Amer. Geophys. Union,1981, p. 215.
17. Molnar, P., *J. Himalayan Geol.*, 1990, **1**, 131.
18. Bilham, R. and Jackson, M., *J. Nepal Geol. Soc.*, 1994, **10**, 18.
19. Chander, R. and Gahalaut, V. K., *Curr. Sci.*, 1994, **67**, 531–534.
20. Khattri, K. N., *Curr. Sci.*, 1993, **64**, 885–888.
21. Gaur, V. K., in *Earthquake Hazard and Large Dams in the Himalaya* (ed. Gaur, V. K.), INTACH, New Delhi, 1993, pp. 63–74.
22. Khattri, K., in *Earthquake Hazard and Large Dams in the Himalaya*, (ed. Gaur, V. K.), INTACH, New Delhi, 1993, pp. 35–62.
23. Gahalaut, V. K. and Chander, R., *Tectonophysics*, 1992, **204**, 163–174.
24. Gahalaut, V. K. and Chander, R., *J. Geol. Soc. Ind.*, 1992, **39**, 61–68.
25. Chander, R., *Tectonophysics*, 1988, **49**, 289–298.
26. Chander, R., *Tectonophysics*, 1989, **170**, 115–123.

ACKNOWLEDGEMENTS. This research was supported by a CSIR–ES scheme and the host institution, the WIHG. I am indebted to P. K. Roy and S. S. Teotia for nudging me into studying earthquake phenomena using the framework of fractals, and also for stimulating discussions.

Received 12 August 1994; revised accepted 5 May 1995

Beyond the Quantum Paradox. Lazar S. Mayants. Taylor and Francis, London. 1994. pp. xi + 105. Price: £9.95.

Among the numerous attempts, of late, to develop a correct point of view in respect of quantum mechanics, this book claims to offer a rather straightforward solution. In doing so some will say it goes a bit too far as, for example, in postulating the existence of emons, i.e. 'concrete' electromagnetic quanta with nonzero mass, even though such ideas have been suggested by others in the past from time to time. To take the last issue first, the case for nonzero mass for the quanta of electromagnetic fields is taken to be based on the formulae[1]

$$p_x = mc\beta_x, \quad E = mc^2, \quad m = m_0\eta^{-1}$$

$$(1 \le \eta^{-1} < \infty), \quad \eta^{-1} = \frac{1}{\sqrt{1-\beta_x^2}},$$

so that if $m_0 = 0$, then $m = 0$ even if $\eta^{-1} \to \infty$ when $\beta \to 1$, and if $m = 0$ then both p_x and E are identically zero. For a nonzero electromagnetic field to exist, m_0 and m must also be nonzero. One may rightly express one's worries here about trying to extract meaning out of expressions like $0 \times \infty$, or 0/0, etc.

The above is, however, only one of the conclusions to which the author is led to, in view of what he calls his principal principle, which is that in dealing with statements involving probabilities and statistics, one should never confuse the 'concrete' object for the abstract object. The simplest way to explain this by the author is to take the case of balls of different colours in a bag. When we talk, he says, of 'a ball in the bag', this is an abstract object which theoretically can possess any one of those different colours, but when we pick a concrete ball out of the bag it has one definite colour only. Statements of probabilities are always about abstract objects, whereas statistical experiments are always done with concrete objects.

Regarding Schrödinger equation, he says, it refers to the situation when the expectation value of the Hamiltonian of the system coincides with the expectation value of the energy of the particle, the wavefunction serving merely as an auxiliary mathematical device for calculating probabilities.

As an illustration of his views, Mayants proceeds to demolish the various quantum-mechanical paradoxes, like Schrödinger's cat paradox or the EPR-Bohm paradox, etc., by simply using his principal principle. In doing this he attempts to emphasize the role of classical concepts even though he does recognize the peculiar role of the superposition principle in calculating quantum-mechanical probabilities. About the Schrödinger's cat paradox, for example, he says that the 'concrete' cat has only one of the two possible values 'dead' or 'alive', of the property 'state of being', whereas an abstract cat does not exist in reality at all. The double-slit experiment is a Gedanken experiment, claims Mayants, because in an actual experiment the separation between slits should only be of the order of a few wavelengths. Bell's theorem, claims the author, has impeccable mathematics, but its premises are wrong. Bell's theorem is not followed in practice because spin values for the same concrete particle for different axes are not permissible. The only thing Bell-like arguments can prove, claims the author, is that quantum physics is incompatible not so much with realism as simply with determinism, which, in any case, being a probability-based theory, is as opposed to a nonprobability-based theory like nonstatistical classical mechanics.

Mayants argues that the situation in the EPR paradox is the result of the law of conservation of momentum and his principal principle, and that there is no action-at-a-distance, no 'knowing' of the result of measurement on one particle by the other. He, in fact, denies any connection between the Bell-inequalities-related experiments and the concrete mechanical system of the EPR, which latter belongs, he claims, to classical mechanics. Regarding Bohm's experiment, he says that it is meaningless to talk of simultaneous values of spin for different axes for a concrete system. Assuming conservation of angular momentum and his principal principle, the experimental outcome of Bohm's experiments is explainable, he claims, without further ado. Similarly, assuming Malus'es law, the principal principle and what is termed as the law of conservation of polarization, the experiments of Aspect *et al.*[2] are also sought to be explained, and it is claimed that there remains no paradox. In dealing with Wheeler's delayed-choice experiment, says the author, the outcome is determined by the last-moment choice, and not by the delay, as the concrete photons can be thought of as being emitted by two different sources, and interference-like effects then arise because of cross terms occurring when squaring the total wavefunction in getting the probabilities.

In contrast to the claims made by Mayants regarding the observance or otherwise of inequalities due to Bell (as also, by implication, due to Clauser–Horne, and others), Pitowsky[3], for instance, goes on to show convincingly that quantum frequencies often violate the facet inequalities of the correlation polytopes. Indeed, Pitowsky claims that violations of the above inequalities by quantum frequencies pose a major problem to all schools of classical probability. Much earlier, Strauss[4] had shown that the algebra of projectors in Hilbert space is only partially Boolean if complementarity in the Copenhagen sense was to be maintained. Even though experiments testing Bell inequalities have more than reaffirmed quantum-mechanical peculiarities, including nonlocalities, Mayants would have none of this. Like Einstein at one time, he would, as stated above, deny even the very existence of uncertainty relations if interpreted in terms of measurement as done by Heisenberg. The author feels strongly that Heisenberg's uncertainty principle deals with abstract quantities and that, therefore, Heisenberg's original interpretation in terms of experiments with concrete objects was wrong. The free particle and the harmonic oscillator cases, he says, violate the principle $\Delta p \cdot \Delta x \ge \hbar/2$, because for the free particle, for instance, $p = \text{constant}$ implies $\Delta p = 0$, giving $\Delta p \cdot \Delta x = 0$ and not $\hbar/2$.

The author does not think it proper to so much as even mention either the deterministic or the stochastic hidden-variable theories, or the many-worlds or many-minds interpretations of quantum mechanics nor does he think it fit even to mention or discuss such important topics as the reduction of the wavefunction, while rejecting nonlocality offhand. The result is that there are strong statements in the book without the provision of adequate justifications.

One is tempted to think that Mayants' plunge for a 'concrete' ontology, and semblance of an attendant realism may have been motivated by the fact that violation of classical constraints like the Bell inequalities can be taken to imply

antirealism, because, like a Humean skeptic, the antirealist may hold that quantum physics does not deal with causes but with phenomena only, and in a world without causes, 'anything goes'. In the event, however, Mayants ends up making sweeping statements which a more careful analysis may show to be wrong.

The chapter entitled 'Personality versus society' of the book seems out of place in what purports to be an account of the way out from the quantum-mechanical paradox, unless the author intended to show that he was demolishing the myths, he saw, of quantum mechanics, in much the same way as he would demolish the myths perpetrated in the name of dialectic materialism and social formation under communism!

All said and done, the principal principle enunciated by Mayants is something to be kept in mind when talking about probabilities and statistics in quantum mechanics, as in any probability-based theory, but, in spite of claims to the contrary by the author, this in itself can hardly be expected to obviate the need for further analysis of the foundational aspects of quantum mechanics.

The final answer, instead of lying with Mayants, may indeed lie somewhere else, based as it would be, on a more careful analysis of issues such as the superposition principle and entanglement, time evolution and reduction of wavefunction, and, of course, nonlocality, and then it would indeed be found, as shown by Redhead[5], that 'quantum mechanics has landed some pretty queer fish'.

1. Mayants, L. S., *Found. Phys.*, 1981, **11**, 577.
2. Aspect A. *et al.*, *Phys. Rev. Lett.*, 1981, **47**, 460; 1982, **48**, 91; 1982, **49**, 1804.
3. Pitowsky, Itamar, *Quantum Probability, Quantum Logic*, Springer, Berlin, 1989.
4. Strauss, M., *Modern Physics and its Philosophy*, Reidel, Dordrecht, 1972.
5. Redhead, Michael, *Incompleteness, Non-locality and Realism*, Clarendon, Oxford, 1990.

HABEEBUL HAQUE ANSARI

Department of Physics,
Aligarh Muslim University,
Aligarh 202 002, India.

The Action Principle in Physics by R. V. Kamat. IPA monographs in physics series. New Age International Publishers Limited, Wiley Eastern Limited, 4835/24 Ansari Nagar, Daryaganj, New Delhi. pp. 201. Price: Rs 400/-.

Given the large number of students which flock to various postgraduate departments in Indian universities every year, the near total lack of books by Indian authors written specially for Indian students, either as basic textbooks or for supplementary or follow-up reading, is quite remarkable. While it is certainly true that most of the students would rather watch TV than study books in their spare time, and bogged down with heavy course loads as they are, certainly not 'outside-the-course' monographs, this is not a sufficient explanation. For, even if only, say, 5% of the students are seriously interested in their subjects, we are still dealing with large numbers. That they do not do so much is partly because reasonably priced books, pegged at an appropriate level, are not easily available. Clearly, here is a Catch-22 situation.

While there are a large number of monographs of good quality published abroad, these are usually not available to Indian students. Most university libraries are in bad shape. The high price of books is a very major deterrent. Also, not many bookstores kept many specialized monographs in stock, and ordering them from publishers can take 6–9 months (who can afford air mail?). It seems quite unreasonable to expect a student to have sufficient patience and sustained interest for so long. 'In the final year, sir, we study other subjects.' The students are like that only.

The Indian Physics Association, by publishing monographs meant for Indian students of postgraduate level, or first year of Ph D is providing a much needed input. The volumes are reasonably priced, and the prices, hopefully, could be brought down further if the series proves popular. It is reasonable to expect that Indian authors would have a much better understanding of mathematical or conceptual difficulties faced by Indian students and, thus, would be better able to adapt the material to our needs. It is an idea whose time has come.

The present volume is by R. V. Kamat, who has been teaching physics students in St. Xavier's College, Bombay, at both undergraduate and postgraduate levels for over 20 years. It deals with the action principle in classical mechanics, electrodynamics, general relativity and quantum mechanics. The book starts with simple problems in calculus of variation, e.g. the famous brachistochrone problem and Fermat's minimum time principle in geometrical optics. These chapters have a large number of problems, of which at least some will be enjoyed by interested readers. This is followed by discussions of Hamilton's principle in classical mechanics. This is a fairly standard fare for M Sc (Physics) students and is dealt with quite satisfactorily here, as also in most standard textbooks on mechanics. The next chapter deals with the action in classical electrodynamics, and derivation of Maxwell equations, starting from the corresponding action.

The author has been careful enough to recaptitulate some of the concepts, and the covariant notation of the special theory of relativity. These are followed by chapters on the action principle for general relativity and for quantum mechanics, however avoiding a full-blooded path integral formulation (that would perhaps require a book on its own). A chapter on action-at-a-distance for much of electrodynamics, which avoids introducing electromagnetic field as a separate entity and defines direct but 'delayed' action-at-a-distance between charges (as proposed by Fokker and Schwartzchild and developed later by Wheeler and Feynman), has been added perhaps as a teaser for students to pursue the subject further.

The book is suitable for M Sc (Physics) or first-year M Phil students. It has a judicious mix of conventional 'textbook material', and somewhat more advanced topics which are usually found only scattered in different textbooks on electrodynamics and relativity of quantum mechanics. The book thus brings out an important unifying principle in physics, in a language accessible to students.

On the debit side, the price of the book at Rs 400 for 200 or so pages is not really low. The number of typographical errors is quite large for a textbook meant for students. It is hoped that

these will be corrected in the next printing. The problems in the latter part of the book are quite routine, of the kind 'verify Eq. 7.8'.

An important topic which has been omitted is the euclidean field theory, which makes contact with the minimization principle of free energy in thermodynamics. I think that it is technically, conceptually and, hence, pedagogically simpler as it deals with real integrals of the type $\int_{-\infty}^{+\infty} \exp(-x^2)\,dx$, and not complex and not absolutely convergent ones like $\int_{-\infty}^{+\infty} \exp(-x^2)\,dx$. Some jargon-free discussions of if and how causality is maintained even with advanced wave solutions in Chapter 8 would have been very useful, and also would have helped the students think about the physical ideas behind the mathematical formalism.

On the whole, the book is a welcome effort to bring good books to Indian physics students at a reasonable price. It is recommended for acquisition by libraries of institutions involved in physics teaching at M Sc or Ph D level, and also by students who want to go a bit beyond the conventional curriculum.

DEEPAK DHAR

Theoretical Physics Group,
Tata Institute of Fundamental Research,
Homi Bhabha Road,
Bombay 400 005, India.

Laboratory Manual of Organic Chemistry. Raj K. Bansal. Third Edition. Wiley Eastern Limited, New Delhi. 1994, 290 pp. Price: Rs 80.

An author who sets out to write a guide on practical organic chemistry for Indian universities is faced with a choice between two evils. Given the backdrop of a tradition-bound, rigid system (wherein innovation is frowned upon), the book has to meet the demands of the existing (though outmoded) syllabi to make commercial sense for both the author and the publisher, or face oblivion. An innovative guide which sets out to correct the course of practical organic chemistry within this country would have few takers. The bulk of the colleges (over 7000) are nothing more than factories churning out graduates within the rigid norms imposed by the examination systems, that establish a uniformly low level by sticking to the letter (not spirit) of the syllabi. The handful of autonomous colleges/institutions who have the resources to modify and innovate either do not have an undergraduate programme or have chartered their own different courses. Writing a book for this miniscule market makes no commercial sense. Unfortunately, what makes commercial sense has very little professional value. Chemistry being a professional science (which has given birth to several professions and even today continues to churn out new professions) should have a dynamic practical course to prepare students to meet the demands of the present job market and future challenges. One would like to see a practical guide which contains not only classical areas, including spectroscopy, but also a good sprinkling of professional skills such as clinical chemistry (biochemistry), pollution chemistry (air and water monitoring), detection of adulterants in food (forensic science) and some basic industrial chemistry. It should set out a model framework for undergraduate (UG) and postgraduate (PG) courses to make the practical courses relevant to the job market. Alas!! an author who sets forth on this trial would have to wait for the world around him to change, for no publisher would support such a venture.

It is against this background that one should view this book. While Bansal has kept in mind the overall needs of the Indian universities, there are some curious features. It is hard to see why any college would like to store 11.5×40 (students) = 460 g of picric acid prepared every year. Several such preparations appear to be there for old times' sake. The analytical organic chemistry experiments baffle this reviewer. Let us assume 14% nitrogen/sulphur/halogen in a given organic compound. If one takes 2 mg of sodium and an equal quantity of the compound (but does not the grand master A. I. Vogel state that the compound should be added to molten sodium) per ignition tube and 15 ml of water, one could hardly expect 0.02 mg of cyanide/sulphide/halide, assuming 50% reaction (which is impossible under the conditions). Should one depend on precipitation tests for sulphide? The existence of sulphocyanide in a very dilute alkaline solution is also questionable (if tests are performed as per procedures on pure organic compounds). In the same way, several other functional group tests are also of little practical value. The heavily curtailed list of common organic compounds indicates clearly that the author wishes to weed out items which are seldom given for UG and PG classes. This laudable objective could have been stretched to other procedures such as estimations. The section on organic preparations has several important and useful procedures for classroom work. Considering the financial crunch, the scales of these preparations are generally large. The section on spectroscopy stands alone without connectivity. The cost of organic analysis could be drastically reduced and the training made more realistic to real-world situations by combining classical and spectroscopic approaches. After N, S, halogen test and m. p. determination, the student could analyse the given (photocopy) infrared (IR) spectrum. The functional groups deduced from IR could be confirmed by chemical tests and derivatization. Listing the derivative under different functional groups could send wrong signal to students. For example, the phenolic group is identified by the $FeCl_3$ test, solubility tests and preparation of a benzoyl derivative. However, an easy derivative for some phenols (and aromatic amines) could be via nuclear bromination. A good derivative for quinol is *p*-quinone, by $FeCl_3$ oxidation. Some glaring errors in figures and structures in this third edition could have been avoided (e.g. water pump (p. 11), thiourea (p. 142), α-bromo-β-napthol (p. 158), piperine (p. 236), etc.).

Notwithstanding some drawbacks, the book meets the demands of most of the syllabi. The discussions and procedural details have been trimmed to a minimum to keep the book affordable to students, who should also find the discussions on mechanisms and chemical equations very useful in understanding the procedures. As frequently suggested in the text, the instructor should provide supplementary guidance, bearing in mind that brevity has its own pitfalls. On reduced scales, the experiments on

named organic reactions would prove to be of good instructional value. Students should find this book affordable and useful for routine laboratory work.

R. BALAJI RAO

Department of Chemistry,
Banaras Hindu University,
Varanasi 221 005, India.

Perspectives in Nematode Physiology and Biochemistry. M. L. Sood and Jyotika Kapur, eds. Narendra Publishing House, 1417 Kishan Dutt Street, Maliwara, Delhi 100 006. 1995. 302 pp. Price: Rs. 575/-, $ 69.75.

Nematodes, much like insects, are one of the most abundant animals, both in their numerical strength as also in the number of species. While insects are very conspicuous by their presence as also because of some 800,000 known species, the knowledge of nematodes with only about 30,000 known species is very scanty because of their hideous and illusive nature, small size, concealed habitats, colourless bodies and the difficulties that are generally encountered in their collection and isolation. In the last 2 to 3 decades, because of the interest that has been generated due to their economic importance, nematology has fast emerged as an important area of parasitology, helminthology, plant nematology as also that of agricultural and biological sciences. There were not many text/reference books on nematodes until recently, more particularly from India. With this background, in my mind, I naturally was very excited by the mere sight of this book, more so because of its title. But my enthusiasm turned into disappointment as I read through the book.

Keeping in view the paucity of information on the physiology and biochemistry of nematodes, one would have expected some well-knit, properly interlinked, and thoroughly edited write-ups, but what we have in its place are some essays on certain aspects of physiology and a couple of them on biochemistry as well. It is interesting to note that two of these write-ups are by scientists who have not even made any contributions, as judged from the bibliography provided, in the area on which they have contributed the articles. Some other topics like 'Test parasites and nematicidal drugs', though of significance and academic interest, do not actually strictly fall under the title of the book. Similarly, the article on *Ascardia galli* appears more to be a Ph D research topic rather than a review article and the same should merit publication in a journal rather than in this book. However, from these remarks it should not be inferred that the book does not have merits and the information contained therein is totally devoid of value and utility. The articles by Bone, Boczone, Srivastava and Batra, Harnett and Parkhouse, etc., are exhaustive, up-to-date and praiseworthy.

I feel that the editors have missed a great opportunity of producing a useful book on such a challenging subject. The choice of topics should have been a more careful one. It also appears that the various contributors did not get proper guidelines from the editors and as a consequence the various write-ups do not appear to be the parts of the same book. It is also strange that no detailed introduction or preface precedes the articles. The quality of printing and the general get-up of the book is satisfactory. There are some typographical and printing errors, which should have been avoided.

M. SHAMIM JAIRAJPURI

Department of Zoology,
Aligarh Muslim University,
Aligarh 202 002, India.

Ocean Technology – Perspectives. Sushil Kumar, Vikram V. Agadi, M. Keshava Das, B. N. Desai, eds. CSIR, New Delhi. 1994. 989 pp.

The International Indian Ocean Expedition of 1960–65 was the first attempt in the country towards systematic studies of the seas around India – culminating in the establishment of the National Institute of Oceanography (NIO) at Goa, which has since been in the forefront of oceanographic research in the country. Besides NIO, research in ocean science and engineering is carried out in more than a dozen research institutes, IITs and universities. A glimpse of the contemporary work in the country in ocean technology is given in the book under review. The book is a compilation of 95 papers presented at a national symposium organized by NIO in August 1992 as a part of the CSIR Golden Jubilee Celebration, to which it is a fitting tribute.

As the title of the book indicates, more than 80 papers deal with different topics in ocean technology and the remainder with ocean science, except for one odd paper on wave propagation in alkaline metals, which is not relevant to the theme of the book. It would have been convenient to the reader, if the editors had grouped the papers under different subheadings. On a study of the book, the following topics can be identified with more than 4 papers in each: ocean waves, wave forces, ocean bed surveys, oil field development, remote sensing, corrosion and biofouling, pollution monitoring and control, biotechnology, deep-sea mining and marine instrumentation. On about a dozen other topics of oceanography, there are one or two papers each. A glaring omission is the subject of deriving electrical energy from the oceans (on which at least three agencies have been working in this country), except for a single paper and that too on the concept of an ocean-coupled solar pond for conversion of thermal energy into electrical power.

There are four papers on marine survey systems and techniques – the very first paper on survey for engineering applications, two papers on the multi-beam sonar for seabed mapping and a review paper on current trends in underwater operations and tools used.

A wide variety of problems related to the design and performance of coastal and offshore structures, guyed towers, tension leg platforms and moored buoys in the presence of ocean waves and currents is discussed in over twenty articles, which will be of immense use to a designer. Four papers which deal with techniques of wave measurements, wave modelling taking into account transformation processes in shallow waters, will be useful in estimating the design wave parameters. A paper on tidal currents in shelf regions deals exhaustively with both theory and practical measurements.

The problem of corrosion in cement, concrete and steel structures in marine environment as well as techniques of studying the corrosion process and corrosion fatigue in metals and biofouling in power plants are discussed in six articles, and will be valuable to engineers.

Out of the twelve papers on marine instrumentation, nine present original work on the development of new instruments, some of which have been field-tested and merit consideration for commercial production, and three give overviews of commercial systems of foreign manufacture and data management systems. The instruments are used for integrated data acquisition and measurement of water quality: spectral irradiance, underwater thermal profile, ambient noise, tidal heights, subsurface currents and core sampling. The article on nondestructive testing is just a review article on current techniques.

An excellent set of eight papers deals with satellite oceanography, namely techniques for deriving from satellite data information on fish abundance zones, littorial sediment transport, sea surface temperature and mixed-layer depth, atmospheric moisture content, surface wind and wave, latent heat flux and bathymetric features. With the advantage of getting synoptic data over wide regions at periodic intervals from the satellites, these papers will be valuable for research studies and applications.

Another topic of contemporary interest, namely, exploration for and exploitation of polymetallic modules in the central Indian Ocean, as well as the environmental protection aspects while mining, is dealt with in about fifteen papers which merit careful study.

Environmental pollution, in particular oil spilling, is the subject of five papers; the use of microbes for oil degradation or recovery is discussed in four papers. Other aspects of biotechnology, such as processing seaweeds for drugs, are discussed in three papers.

Among the fifteen papers dealing with different aspects of ocean science, some are significant and worth follow-up, such as the spatiotemporal variability of biological and physical fields, simulation of acoustic propagation, storm surge modelling, classification of acoustic echoes from seabed with broadband signals and directional response computation for an arbitrary acoustical array for broadband signals.

The usefulness of the book would have been enhanced by introducing each of the ten major topics with a summary of the present status as revealed in the papers and the perspectives for future. A symposium of such a high order as the one under review would have seen valuable discussions at the end of presentation of each paper and their inclusion would have further enhanced the value of the book. However, even in its present form, the book is highly informative and will be valuable to any research worker in emerging fields of ocean technology.

D. SRINIVASAN

Emeritus Scientist,
Camp NIO,
Goa 403 004, India.

Botany in India: History and Progress. Vol. I. B. M. Johri, ed. Oxford & IBH Publishing Co. Pvt. Ltd., 66 Janpath, New Delhi 110 001. 1995. xxxi + 521 pp. Price: Rs 950.

The above publication, a multiauthored volume, comprises consolidated information primarily on lower groups of plants from India. During the past three decades or so, there has been phenomenal increase in research activities in the country not only on fundamental aspects of botany but also on several new disciplines/aspects that have been initiated and fruitful research conducted on them. The volume of research publications has been enormous and it is difficult to keep abreast of the information on all the facets of one's own speciality, leave aside botany as a whole. Almost all the contributing authors (except probably one) of the above volume are well-established specialists in their own fields of research and have, therefore, done a fair job of assembling and arranging the research information under neatly segregated sections and subsections. One is often tempted to be subjective in such write-ups, but it is creditable that invariably almost all chapters have an objective approach in their contents.

Historical accounts for the different groups of plants of India have been published earlier at intervals, but probably reviewing them at the beginning of each chapter was evidently a necessity for understanding the continuity of research progress of the subject matter in the subsequent period of time. This has in some cases entailed a partial repetition in some chapters (on fungi).

The title of this publication, in a strict sense, indicates that it gives information only on botanical research carried out in India on Indian plants, by Indian and foreign nationals. Exotic plants studied in India, such as *Tmesipteris* spp. by Sahni at Lucknow (p. 456), could also come in this category. The studies on Indian plants carried out in foreign countries by Indian or foreign nationals, such as *Flora of British India* by J. D. Hooker (p. 22) worked out at Kew (this fact is withheld), actually represent 'botany *of* India', and such instances get confused or even synonymized as 'botany in India'. However, research on exotic plants carried out outside India by Indian nationals is *certainly not* 'Botany in India', and the editor has been irresolute in allowing such information in the publication to persist, (examples given later).

Since an up-to-date consolidation of information on different branches/aspects of botany in India was not available at one place, the above publication is a welcome venture, which will positively provide referential information to postgraduate students of botany for competitive tests, selection of research topics/disciplines/branches to be pursued at particular centres of that speciality in the country. It may, however, be mentioned that the university botany departments are dynamic as far as the specialities are concerned, because they change with the change/retirement of the personnel – a universal phenomenon, and the specialities in the universities may not necessarily be existent.

The matter dealt with in the different chapters is of high standard and detailed, certain chapters being *par excellence,* having been prepared by specialists with large number of publications to their credit. A major part of the book (Chapters 4–24) deals with the lower groups of plants (viruses, bacteria, algae, fungi, lichens, bryophytes and pteridophytes – some of which were earlier known as cryptogams).

An exploratory history and publication of numerous floras (entirely angiospermic) along with the institutions/agencies known for angiosperm taxonomic work have been outlined in Chapter 2. A much detailed and interesting exploratory history by K. Biswas (in 1943) has apparently not been consulted as there is no citation under the refer-

ences. Information about the number and the presence of the total holotypes/isotypes of Indian angiosperms preserved in Indian herbaria, a must for taxonomical work, is lacking. Information on plants for medicine (Chapter 3), primarily angiospermic in content, could have been more comprehensive, and would probably have been appropriately included in Volume II of Botany in India, where several other aspects of angiospermic research have been envisaged to be published.

In Chapters 4–24 (except 20) the approach to each group of plants is somewhat on similar lines in respect of historical account, structural morphology, taxonomy, physiology, genetics, ultrastructure, etc. Detailed information on the characteristics of the individual groups is well treated. Unfortunately, Chapter 20 on lichens reveals lack of understanding of the concept of lichens and lichen components ('... dearth of information on the mycobionts of Indian lichens', fruit body 'perithecium in *Cetraria*' – an absurdity), inaccurate statements ('lichen research initiated at the University of Lucknow by Biswas'), prominence to the work on foreign lichens outside India, and several misspelt terms and lichen names.

The following shortcomings/inaccuracies need reporting so that errata may be published if thought proper:

1. Citation of plant species under abbreviated generic name without the full generic name *immediately* preceding, and if preceded intervened by other genus or genera is liable to cause confusion if the first letter of the genera is the same.

2. Inaccurate statements (e.g. p. 456): 'Sahni (1915, 1916), at Lucknow, ...'. Sahni was at Cambridge University during those years. He joined Lucknow University in 1921.

3. Matter not eligible for inclusion in this book:

(i) Sinha and Ashworth (p. 237) on genetic recombination of *Dictyostelium discoideum.*

(ii) Cooney and Emerson (p. 243) on thermophilic fungi.

(iii) Garrison *et al.* (p. 243) on thermostability.

(iv) Schmitz and Srivastava (p. 204) on ultrastructure of *Laminaria groenlandica*, etc.

(v) LaRue and Narayanswami (p. 437, 440) on *Lunularia cruciata.*

(vi) Rai (p. 381) on carbon–nitrogen ratio on *Peltigera aphthosa, Platismatia* (misspelt Platimatia) *glauca* and *Pseudevernia* (misspelt Pseudoevernia) *furfuracea* (= *Evernia furfuracea*). None of the taxa is an Indian lichen.

(vii) Le Blanc and D. N. Rao and Le Blanc *et al.* (p. 382): work was done on Canadian lichens in relation to pollution in Canada.

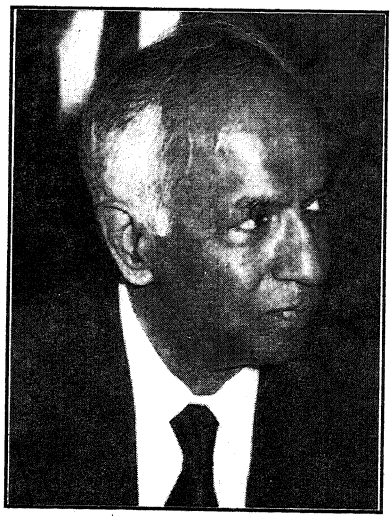
There are a large number of misspelt words, including botanical names of plants; some of them may be due to the printer's devil, but others do not seem to be so.

The price is prohibitive.

D. D. AWASTHI

38 Khurshed Bagh,
Lucknow 226 004, India.

Subrahmanyan Chandrasekhar



1910–1995

The news of Subrahmanyan Chandrasekhar's death on 21 August 1995 arrived as this issue of *Current Science* was going to the press. Chandrasekhar's passing marks the end of a glorious chapter in modern physics. Awarded the 1983 Nobel Prize, Chandrasekhar was one of the dominant figures of twentieth century astrophysics. A forthcoming issue of this journal will carry a detailed account of his scientific accomplishments.

Professor T. R. G.: An appreciation

T. R. Govindachari will shortly be turning 81 and would have seen a thousand moons. I had the privilege of being associated with him for over four-and-a-half decades, as a student in his intercollegiate lectures for B Sc (Honours) in chemistry (long since discontinued), as his first Ph D student and as a CSIR postdoctoral fellow at Presidency College, Madras, as a Departmental Head during his direction of CIBA Research Centre, Bombay, and then on as a fellow organic chemist. I meet him off and on in some symposium or the other, and when I am in Madras, I unfailingly visit him in the compact but well-equipped Centre for Agrochemical Research which he has organized for SPIC Science Foundation. These visits are invariably inspiring and I return wondering whether the country cannot find in his personality panaceas for many an ailment that afflict Indian science.

Some weeks ago I telephoned to him to say that I would be in Madras for a couple of days. He immediately insisted that I should visit him for discussing the affairs of the National Organic Symposium Trust (NOST), a nonprofit organization established by five of the country's seniormost organic chemists, S. Swaminathan, Nityanand, Sukh Dev, Bhattacharya and himself. The major activity of NOST is to promote Organic Chemistry in India by holding conferences (like Guha and Gordon) and to support other symposia from funds that he and the other founders had collected. He had come to Bombay in 1987 for a fund collection drive. He was then 72 and I still remember vividly how he went from one industrial house to another soliciting donations. His strenuous efforts and those of other founders finally gave NOST the income to organize the NOST conferences, which are done by the Council, with the Trustees (TRG being one of them) keeping themselves in the background. NOST is carrying on its mission quietly and with very little fanfare, but, like other acts of good faith, has some disbelievers. TRG shared his concerns with me during this last visit and discussed with me at length how real or perceived imperfections in the organization could be addressed.

Actually, when I stepped into his laboratory that day, I found him engrossed in deciphering an NMR spectrum along with an associate. He told me with his usual self-deprecatory smile how he was trying to inform himself of recent techniques in NMR spectroscopy so that these could be mobilized for structure elucidations. He also divulged to me with pride and satisfaction that he was able to persuade SPIC and DBT to fund a modest Supercon NMR for the Research Centre, and with mischief in his eyes, he showed me how he had converted one floor into two by building a mezzanine, thus effectively doubling the laboratory space, enabling both the research group and the sizeable array of sophisticated equipment to co-exist comfortably. True to his colours, TRG at 81 goes about his research activities, quietly keeping abreast with recent scientific developments and with the tools available to him, not cribbing at the lack of one or the other but silently working towards adding the necessary ones.

On an earlier visit, he had shown me with undisguised excitement, beautiful crystals of azadirachtin A (the now universally famous antifeedant principle present in the seed kernels of neem) and the results of single-crystal X-ray studies (*Current Science*, 1994, **66**, 295, 362). Formulations of crude neem extract containing 300–3000 ppm of azadirachtin A are marketed now in many countries as an eco-friendly pesticide. TRG's group has developed processes for obtaining azadirachtin-enriched neem extracts as well as formulations which are now distributed by SPIC Biotech. I would not be surprised if he had the purest sample of azadirachtin A in the world. TRG also mentioned to me how he had a few other leads from plant sources with pesticidal properties. In fact, for nearly six decades, TRG has brought in unparalleled perseverance, single-minded devotion and incisive chemical insight to the study of the vast flora of this country.

Over the years, his group has isolated many complex-structured compounds belonging to diverse classes of natural products such as alkaloids, oxygen heterocycles and terpenes, quite a few of them being the prototypes of a particular group, e.g. tylophorine, echitamine, kopsine, ancistrocladine (alkaloids) wedelolactone (furocoumarin) and cedrelone and ishwarone (terpenes). Some of them had potent biological activity (tylophorine – antiasthmatic; wedelolactone – hepatoprotective, ancistrocladine – potential anti-HIV). TRG and coworkers had located a very rich source of the anticancer alkaloid, camptothecin in an Indian plant, when only milligrams were available for clinical trials in USA. Currently, TRG's extensive work on neem has yielded many new congeners of azadirachtin A.

I have dealt at some length on TRG's contributions to natural-products chemistry because I feel that our young organic chemists can learn a lesson from it. Currently, there is a resurgence in this branch of chemistry, particularly as



a source of ecocompatible bioactive molecules serving as leads or ending up as drugs or plant protection chemicals. Taxol, artemisinin and azadirachtin are but three recent examples. Let us ask ourselves the following question: Can we do interesting and useful chemistry in this area doing justice to our plant wealth, or rather, should we pursue the fashions of the day in vogue elsewhere. There is no doubt that science is universal and knows no barriers, but why cannot we be trendsetters for a change?

An X-ray crystallographer of the Madras University Biophysics Department told me the other day that even now TRG himself comes to the department personally with samples for structure determination. So does he for special NMR and mass spectra. It is typical of him even past 80 to be in the laboratory every morning regularly at 9 a.m. and work till late in the evening – surely a trait worthy of emulation by the younger generation. His achievements after retirement from CIBA Research Centre in 1975, such as establishing two research centres, speak for his indefatigable energy and passion for science. Besides development of useful processes for synthetic molecules and extraction of natural products, TRG has added about 30 papers to his earlier tally of 274.

Looking back, TRG can be justifiably satisfied with what he has achieved during 45 years of professional career. As a Professor of Chemistry, Presidency College, Madras (1950–1962), he built up a school of research in natural-products chemistry (and that too with minimum available facilities) whose excellence has been acknowledged internationally. Simultaneously, he was acclaimed as a good teacher who ushered in an appreciation of modern theories of organic chemistry. For a year and a half (1961–1962), he also served as the Principal of Presidency college and might have well become the Director of Public Instruction. The next twelve-and-a-half years (1963–1975) saw him as the Director of Ciba Research Centre, which he moulded into an internationally known research centre and wherein he collected a group of the finest young scientists of the country to do basic research for the development of new drugs and dyestuffs. For a couple of years afterwards (1976–1978) he was a consultant to the Central Leather Research Institute, Madras. Later on he established a good R&D facility for Amruthanjan (1977–1986). He joined SPIC in 1987 as an adviser to their Science Foundation and set up the Centre for Agrochemical Research, wherein he continues his activities. In both these laboratories, TRG has been investigating natural products, in addition to guiding the groups in process development activities.

TRG's achievements have earned him several national recognitions (Bhatnagar Prize for Chemistry, Meghnad Saha Medal, Golden Jubilee Commemoration Medal, Fellowships of IASc and INSA) and international honours (Plenary lecture at the first IUPAC, National Products Symposium, Membership of the Bureau of IUPAC) but I firmly believe that many more should have accrued to him. He has been a President of the Indian Chemical Society but not of the prestigious Science Academies. I suspect that his long post-academic association with industry, especially with a multinational company, has been responsible for his name not appearing in the Indian Government's list of New Year Honours year after year. Shy by nature and reticent by disposition, TRG is not gregarious or garrulous and certainly not given to beating his own drums. It is a pity that his experience and expertise in matters pertaining to education and academic and industrial research, especially in new drug development, have not been adequately appreciated and sufficiently utilized by the scientific community. The loss is not his but of Indian science and that is another ailment that afflicts Indian science!

K. NAGARAJAN

Recon R&D Centre,
32/1, Kalena Agrahara,
Hulimavu Village Panchayat,
Begur, Hobli,
Bangalore 560 083, India.

MEETINGS/SYMPOSIA/SEMINARS

Short Lecture Courses in Neurobiology

Date: 16–22 October, 27 November–6 December and 6–13 December 1995
Place: Bombay
Date: 26 October–2 November, 7–14 December and 26 November–5 December 1995
Place: Bangalore

Topics included for respective dates and places are: Synaptic transmission, Vesicular events at the synapse and neural development.
Candidates should apply to the course location that is closer to them.

Contact: Dr K. S. Krishnan
Molecular Biology Unit
TIFR, Bombay 400 005
and
Dr M. K. Mathew
NCBS, TIFR Centre
P.O Box 1234, Bangalore 560 012

Third International Symposium on Genetics, Health and Disease

Date: 1–4 December 1995
Place: Amritsar

Topics include: Genetic diseases; Diagnosis; Treatment and counselling.

Contact: Dr Jai Rup Singh
Centre for Genetic Disorders
Department of Human Genetics
Guru Nanak Dev University
Amritsar 143 005
Phone: 91(183) 258863
Fax: 91 (183) 258820

Eugene Paul Wigner – A tribute*

N. Mukunda

One of our last surviving links with the period of the creation and development of quantum mechanics was broken with the passing of Eugene Wigner on 1 January 1995 at Princeton in the USA. Wigner was remarkably talented and wide-ranging in his interests, and his work touched innumerable aspects of modern physics. In every area that he turned to, he discovered new and profound insights and interesting viewpoints, often understood and carried further by others much later. He was as much at home in fundamental problems of physics and their mathematical analysis as in engineering and technological matters. In this tribute, I shall first describe briefly his life and career, then turn to a sketch of his work, and conclude with an attempt to capture his personality and philosophy of science and life.

A brief life sketch[1]

Eugene Paul (Jeno Pal in Hungarian) Wigner was born on 17 November 1902 in Budapest, Hungary, to Elisabeth Einhorn and Anthony Wigner. He thus belonged to the same generation as Werner Heisenberg, Enrico Fermi and Paul Dirac. Leo Szilard and John von Neumann were Wigner's classmates at the Lutheran High School in Budapest – 'at that time, perhaps the best high school of Hungary and probably also one of the best of the world'[2]. Wigner retained great regard for his mathematics teacher L. Ratz, who also recognized and encouraged von Neumann's unusual talents.

After a year spent at the Technical Institute in Budapest, in 1921 Wigner joined the Technische Hochschule in Berlin to train as a chemical engineer. He completed his doctorate in 1925 and then worked for a year and a half as a leather chemist. By this time he had become very much a part of the Berlin physics scene; his break came with an appointment as assistant to Erwin Schrödinger (who had succeeded Max Planck at Berlin) for 1926–27. This was followed in 1927–28 by a position as assistant to Max Born in Göttingen, and then as Privatdozent at Göttingen during 1928–30. At this point he moved to the United States, where he spent the rest of his life.

Wigner's career in the US began as a lecturer in mathematical physics during 1930 at Princeton University, quickly elevated to a Professorship from 1930 to 1936. The year 1937–38 was spent as a professor at the University of Wisconsin at Madison. Upon return to Princeton, he became Thomas D. Jones Professor of Mathematical Physics in 1938, a position he held until 1971. The academic year 1957–58 was spent at Leiden in the Netherlands.



In 1937 Wigner became a naturalized citizen of the United States. He took his citizenship very seriously, and played a very active role in public affairs and matters of government policy. As his contribution to the war effort, he spent 1942–45 at the Metallurgical Laboratory of the University of Chicago, the last two years as the head of the theory group there. Earlier he had joined Szilard and Fermi in persuading Einstein to write the famous August 1939 letter to President Franklin Roosevelt that led to the setting up of the Manhattan Project. He was present at the University of Chicago's Stagg Field Squash Courts on 2 December 1942 to witness the world's first controlled nuclear fission reaction set up under Fermi's leadership. During 1946–47 he served as Director of what later became the Oak Ridge National Laboratory in Tennessee. In 1952 he was full-time adviser to the Du Pont Company to design the Savannah River heavy-water plutonium production reactors. Soon after, in 1954 he was appointed to the General Advisory Committee of the United States Atomic Energy Commission, and served also on many panels of the Science Advisory Committee to the President of the United States.

Of the many awards that came to Wigner, we must mention the Medal for Merit, the Franklin Medal for 1950, the Enrico Fermi Award of the USAEC for 1958, the Atoms for Peace Award for 1960, the Max Planck Medal for 1961, and the 1963 Nobel Prize in physics (shared with Maria Mayer and Hans Jensen) for his wide range of contributions to quantum mechanics.

Wigner's first marriage, to Amelia Franck in 1936, was followed by a second one in 1941 to Mary Annette Wheeler, a professor of physics. His sister Margit Balasz nee Wigner married Paul Dirac in 1937. It appears that Dirac was so shy that he once introduced his wife to an old friend as Wigner's sister. In response (!) Wigner referred to Dirac as 'my famous brother-in-law'[3]. There is a charming account by Margit of her first meeting with Dirac in Wigner's company. At a meeting in Budapest, the von Neumanns had invited Margit to visit and stay with them in Princeton. And then[4]:

'Eugene insisted, "If you come to Princeton, you must stay with me. What would people say if you did not stay with your brother?" I was not terribly thrilled with the idea. The von Neumanns had a lovely home,..., while my brother liked to appear, and act, like a pauper. We sailed in the fall; Eugene had a two-bedroom apartment, proudly boasting that he furnished it to the cost of under $25. It looked like it ... It was soon after my arrival; we were having lunch at one of these restaurants, when a tall, slender young man entered the dining room, looked at Eugene and

*Text of special colloquium given at Raman Research Institute, Bangalore on June 16, 1995.

greeted him. He looked lost, and sad. I asked who he was, still standing undecided and none too happy-looking. I was told, he was an English physicist, whom Eugene knew in Göttingen, where they used to have their meals together. "He does not like to eat alone". "So why don't you ask him to join us?" That was how I met Paul Dirac. That was the fall of 1934. The Institute for Advanced Studies had no building of its own as yet. Its members, like Einstein, von Neumann and Dirac as a visiting member, had adjoining rooms in a large university building, called Fine Hall. I remember so well: to the left was Einstein's room, in the middle Eugene's and to the right of him, Dirac's.'

Wigner, Szilard and von Neumann formed the famous Hungarian trio who contributed so decisively to intellectual life in the United States in the 1930s and later. There is a story that during a meeting of scientists connected with the war effort there was so much confusion due to many languages being used that someone got up and exclaimed: 'Gentlemen, let us use one language we can all understand – Hungarian!'

When Wigner died he left behind his third wife Eileen, a son and two daughters.

Contributions to science and engineering

Wigner's work in physics is characterized by hard mathematical analysis based on simple yet profound physical assumptions. While there is a down-to-earth practical quality to some of his work, in others he dealt with the most fundamental issues with great refinement – he was both an artist and an engineer, and quantum mechanics was his medium. To quote John A. Wheeler[5]: 'In the work of Eugene Wigner one sees the basic harmony between the conceptual framework of physics and the structure of the mathematics associated with that physics.' On the other hand, his grasp of technology is best conveyed by this passage from Lawrence Dresner and Alvin M. Weinberg[6]: '.... the facility with which he could pass back and forth between engineering and physics – from a discussion of the probable distribution of energy levels in U^{235} to a critical examination of the blueprints of the concrete foundations for the Hanford reactors, or from a group theoretical argument in transport theory to the design of aluminium fuel elements!'

Wigner's first important work in physics, completed during his apprenticeship with Schrödinger, was a powerful treatment of quantum many-fermion systems. Around the time of the move to Göttingen, and following a suggestion by von Neumann, he undertook the major task of introducing group-theoretical methods into quantum mechanics. By 1928 he had published six landmark papers on the subject; he shares with Hermann Weyl the credit for making this an essential and characteristic component of quantum physics which pervades all its applications. During the 1930s he worked in solid-state physics and at the frontiers of the developing subject of nuclear physics, making a major effort to understand the forces between nucleons, and developing the compound nucleus model to explain resonance phenomena in neutron-induced nuclear reactions. His development later of the R-matrix theory of nuclear reactions was a response to a comment by Fermi that the compound nucleus model needed a firm theoretical foundation. Probably his most remarkable work in mathematical physics – the study of the unitary representations of the inhomogeneous Lorentz group – grew out of a suggestion made to him by Dirac in 1928. This was completed in Madison in 1937, and subsequently became the basic framework for all relativistic quantum theories. He came back to problems of nuclear structure in his supermultiplet theory of 1937, and later in his statistical treatment of nuclear spectroscopy based on random matrices.

In the midst of all this, in the 1940s he worked on the theory of neutron chain reactors and the design of plutonium breeder reactors.

Wigner's concern with the structure of quantum mechanics has led to a series of incisive insights over many years. In the early 1960s he turned to problems of interpretation and epistemology raised by the standard interpretation of quantum mechanics. At this point it is convenient to present briefly and selectively sketches of Wigner's work under several broad areas. This is admittedly an inadequate, incomplete and possibly superficial way to survey his work, yet it may succeed in conveying some idea of the range and magnitude of his achievements. Before embarking on this, we may recall the following important books published by Wigner: (1) *Group Theory and its Application to the Quantum Mechanics of Atomic Spectra* (Academic Press, New York, 1959; the original German version published by Friedrich Vieweg, Braunschweig, 1931); (2) *Nuclear Structure*, with Leonard Eisenbud (Princeton University Press, 1958); (3) *The Physical Theory of Neutron Chain Reactors*, with Alvin M. Weinberg (University of Chicago Press, 1958); and (4) *Symmetries and Reflections – Scientific Essays* (Indiana University Press, 1967). We may also mention that the October 1962 issue of the *Reviews of Modern Physics*, published on the occasion of his 60th birthday, contains many articles surveying Wigner's work in several areas.

Structure and content of quantum mechanics

Any serious user of quantum mechanics is sure to find herself employing repeatedly, either explicitly or implicitly, one or another of the many basic concepts and methods invented by Wigner. One of the earliest is the concept of parity[7]. In classical physics, space inversion is merely a geometrical operation or transformation, a rule to map each point in space to its image by inversion through a chosen origin. The time is left unaffected. A particle trajectory, for example, would be mapped on to another possible trajectory.

Classical space inversion:

$$P: x \to -x, t \to t,$$

$$x(t) \to -x(t).$$

Wigner showed that in quantum mechanics, parity is more than a transformation, it is a physical observable whose value can be experimentally measured. The possible results of measurement are ± 1, and the corresponding quantum states are said to possess even or odd parity, respectively.

Quantum space inversion:

$$P\psi(x,t) = \psi(-x,t),$$

$$\psi(-x,t) = \pm\psi(x,t) \Rightarrow P = \pm 1,$$

even/odd states.

It was this role of parity in quantum mechanics that was shown by Wigner to be the explanation for Laporte's selec-

tion rule in atomic spectroscopy[8]: the matrix elements of the electric dipole moment operator, and hence the corresponding transitions, vanish unless the two concerned states have opposite parities.

The deep connection between invariance principles and conservation laws, both in classical physics and in the quantum domain with specifically new and subtle features, remained a lifelong concern for Wigner, something he came back to time and again. In the particular case of rotational symmetry, the general programme of incorporating group-theoretical methods into quantum mechanics led to Wigner's impressive body of results concerning angular momentum in quantum mechanics[9]. The detailed representation theory of the rotation group SO(3) and its covering group SU(2), which is basic to quantum mechanics, was developed by him in a form suited to practical application. The angular momentum addition theorem, the concept of tensor operators, the Wigner–Eckart theorem for their matrix elements, explicit expressions for the Clebsch–Gordan coupling coefficients (also called the Wigner $3j$ symbols), leading on to the intricate Racah–Wigner calculus for coupling of tensor operators and computing the resulting matrix elements, the generalizations to other symmetry groups – all these oh-so familiar tools of the trade in atomic, nuclear and particle physics originate from his work.

In his book on group theory, Wigner formulated and proved a fundamental theorem concerning the representation of symmetry operations in quantum mechanics[10]. This is a very deep and subtle result, and a brief explanation would not be out of place. The relation between physical states and wave functions (or Hilbert space state vectors) in quantum mechanics is one-to-many. This is because a change in the overall phase of a wave function is physically unobservable:

$$\text{vectors } \psi, e^{i\alpha}\psi, e^{i\beta}\psi, \ldots \rightarrow \text{same physical state,}$$

$$\text{vectors} \xleftrightarrow{\text{Many-to-one}} \text{physical states.}$$

Therefore, what physical states correspond to in a one-to-one manner are not vectors but rays: a ray is an equivalence class of vectors, two vectors being declared equivalent if they differ only by a phase. The ray to which a vector ψ belongs can be unambiguously described by the corresponding projection operator or density matrix ρ_ψ:

$$\text{vector } \psi \rightarrow \text{ray } \rho_\psi = \psi\psi^+,$$

$$\text{rays} \xleftrightarrow{\text{one-to-one}} \text{physical states.}$$

Rays do not form a vector space, so their geometry is somewhat harder to visualize than that of vectors ψ. Wigner's theorem then shows that any mapping $\mathcal{T}$ of rays (i.e. physical states) onto themselves preserving quantum-mechanical probabilities – and any symmetry operation must be of such a nature – can be 'lifted' to either a linear unitary or an antilinear unitary (antiunitary) transformation T on vectors.

Unitary–antiunitary theorem:

Symmetry operation $\mathcal{T}$,

$\rho_\psi = \psi\psi^\dagger \rightarrow \mathcal{T}(\rho_\psi) = \rho_{\psi'} = \psi'\psi'^\dagger$

$\rho_\varphi = \varphi\varphi^\dagger \rightarrow \mathcal{T}(\rho_\varphi) = \rho_{\varphi'} = \varphi'\varphi'^\dagger$

$\psi', \varphi', \ldots$ determined up to phases,

$|(\varphi', \psi')| = |(\varphi, \psi)| \Rightarrow$

either

$\psi' = T\psi, \varphi' = T\varphi, \ldots,$

T linear unitary,

$(\varphi', \psi') = (\varphi, \psi) \rightarrow$ unitary alternative;

or

$\psi' = T\psi, \varphi' = T\varphi, \ldots,$

T antilinear unitary,

$(\varphi', \psi') = (\varphi, \psi)^*$

$= (\psi, \varphi) \rightarrow$ antiunitary alternative.

(Here the inner product of the Hilbert space vectors φ, ψ is denoted by (φ, ψ)). This remarkable theorem has been extended and proved under different conditions by others over the decades.

Most symmetries in quantum mechanics turn out to be of the unitary type; time reversal is one example where the antiunitary alternative is realized. The analysis of this transformation in quantum mechanics was given by Wigner[11] in 1932. In Schrödinger's quantum mechanics, time reversal acts on wave functions thus:

$$T\psi(x, t) = \psi(x, -t)^*.$$

Unlike parity, however, this operation does not have the status of a physical observable in quantum mechanics, and its eigenvalues are not invariantly defined and are not experimentally measurable.

Continuing with the theme of symmetry in quantum mechanics, Wigner and von Neumann proved a very interesting result in 1929, which is of great importance especially in molecular physics[12]: if the electronic states in a molecule are classified according to their symmetry, i.e. the representation of the full group of symmetry of the relevant molecule, and if we have two distinct eigenvalues and eigenstates sharing the same symmetry (two electron terms), these eigenvalues will not cross (become accidentally equal) as one varies the internuclear distances in the molecule. On the other hand, electron terms of distinct symmetry can cross. This is a general theorem of quantum mechanics, applicable to a generic hamiltonian possessing some symmetries and dependent on a continuous parameter: as the parameter is varied, distinct eigenvalues 'of the same symmetry' will not accidentally cross but will repel each other.

Many years later, Wick, Wightman and Wigner[13] brought to light another aspect of symmetry in quantum mechanics, namely the existence of superselection rules. This amounts to a restriction on the applicability of the superposition principle in quantum mechanics. In general, the Hilbert space of states of a quantum system breaks up into sectors, and the formation of complex linear combinations to produce new states from old is limited to one sector at a time, not cutting across sectors. This is the reason why the phase of a spinor field – a field with half odd integer spin – is nonobservable. So, for instance a nontrivial linear combination of states with integer and half odd integer angular momenta cannot be prepared. As another example, one finds that linear superpositions of states of distinct electric charge are unphysical. It is suspected that these results had long been known to Wigner, and he was persuaded by his coauthors to join them and say so in print.

In the preface to his book on group theory, Wigner relates a conversation with von Laue on the use of group theory as the natural tool with which to tackle problems in quantum mechanics[14]. He says: 'I like to recall his question as to which results ... I considered most important. My answer was that the explanation of Laporte's rule (the concept of parity) and the quantum theory of the vector addition model appeared to

me most significant. Since that time, I have come to agree with his answer that the recognition that almost all rules of spectroscopy follow from the symmetry of the problem is the most remarkable result'.

The exponential decay law for unstable states has been well known since the days of Rutherford's experiments on radioactivity. The first properly quantum-mechanical discussion and derivation of this law is due to Weisskopf and Wigner[15]. They were able to provide the basic theory for the natural linewidths and lifetimes of atomic states decaying via transitions to other states with emission of radiation. Their use of second-order perturbation theory along with judicious and delicate assumptions also disclosed that the exponential decay law is only an approximate, not an exact, consequence of quantum mechanics; so departures from it for both very short and very long times are to be expected.

The linear superposition principle of quantum mechanics, already referred to above, finds its most natural expression at the level of state vectors in Hilbert space. The ray space or density matrix description of physical states, which is closer to a classical description, obscures this principle somewhat – it is present but not manifest. In 1932, while studying thermodynamic equilibrium in quantum mechanics, Wigner introduced another description of states for quantum systems possessing classical canonical analogues[16]. Thus, each quantum state is describable by a certain real distribution or function on the classical phase space. In one dimension with classical phase space variables x and p, the construction is as follows:

$$\psi(x) \to W(x,p) = \frac{1}{h}\int_{-\infty}^{\infty} \mathrm{d}x' \; \psi\left(x-\frac{1}{2}x'\right)\psi\left(x+\frac{1}{2}x'\right)^{*} \exp(\mathrm{i}x'p/\hbar)$$

These distributions – named after Wigner – are at the level of density matrices, not state vectors. They are suggestively like classical probability distributions on phase space, such as one uses in classical statistical mechanics. However, since in general $W(x, p)$ can become negative for some arguments, we do not have a classical statistical picture with well-defined probabilities. This is as it should be, since quantum features must be preserved. This description of states in quantum mechanics turned out to be the counter-part, or companion, to a rule or convention given by Weyl for associating a quantum-mechanical operator with every classical dynamical variable; and these ideas were further extended, particularly by Moyal[17].

Wigner contributed a great deal to the formal description of scattering and reaction processes in quantum mechanics, especially in the context of nuclear physics. One of his results concerns the physical meaning of phase shifts. In general, scattering cross-sections are determined by the squared magnitudes of S-matrix elements, and in these the phases get washed out. On the other hand, the spatiotemporal development of a scattering process described within the limits set by quantum mechanics, involves these phases. The beautiful connection found by Wigner is the expression for time delay caused by interaction and its relation to the energy dependence of the scattering phase shift[18]:

$$\Delta T(E) = 2\frac{\mathrm{d}}{\mathrm{d}E}\delta(E)$$

Here $\delta(E)$ is the phase shift at energy E; thus, an attractive (repulsive) interaction leads to $\delta(E)$ increasing (decreasing) with energy, hence to a delay (advance) in the appearance of the final-state products of a collision after undergoing interaction.

We conclude this account with a couple of 'curios'. Classically, one expects the possible states of a system of interacting particles – especially, a two-body system – to separate into two types: unbounded or scattering states, having positive energy, and bound states, having negative energy. In quantum mechanics we expect the energy eigenvalues to behave analogously: a continuum of unbound, positive-energy scattering states sitting on top of a set of discrete negative-energy, bound states. Only the latter have normalizable wave functions. In a remarkable paper in 1929, Wigner and von Neumann produced an example of a two-body potential which possesses a bound state embedded in the continuum[19]! This is an unexpected and essentially quantum-mechanical result. The potential is 'artificial' in that it has to be carefully engineered to produce the desired result, and the state involved is unstable even under small perturbations.

The passage from classical to quantum mechanics results, at the level of dynamical variables, in the loss of *commutativity* in multiplication. Thus, for two physical quantities represented by operators A and B, in general $AB \neq BA$. However, this departure from the classical is limited in the sense that *associativity* is maintained: for three (or more) quantities multiplied in a given sequence the product is unambiguous: $(AB)C = A(BC) = ABC$. One can ask how quantum mechanics might be modified if one takes the nonclassical path one step further and, along with commutativity, one gives up associativity as well. This was examined by Jordan, von Neumann and Wigner[20] in 1934 . It did not, however, lead to any alternatives with sufficiently interesting and flexible properties to give a further extension of quantum mechanics.

Going over this rich list of contributions, one is tempted to say that Wigner took his revenge for not having been involved in the discovery of quantum mechanics, and compensated for it accordingly!

Nuclear forces, structure and reactions

Following the discovery of the neutron by Chadwick in 1932, there was a great deal of work exploring the nature of the strong nuclear forces between neutrons and protons. It was realized that these would be strikingly different from the familiar Coulomb forces between protons, of very short range, and with complicated distance dependences. Further dependences on spin and space exchange were also anticipated. Wigner was one of the earliest contributors to this field, and his name is associated with one of the four basic types of terms in the potential energy expression[21]:

potential energy between proton and neutron =

purely distance-dependent Wigner term +
spin exchange Bartlett term +
space exchange Majorana term +
spin and space exchange Heisenberg term.

Thus, the Wigner force is the simplest of all; the others either distinguish between singlet and triplet spin states, or between even and odd orbital angular momenta, or both. Such phenomenological potentials are useful in analysing low-energy nuclear bound states, scattering processes, etc.

The low-energy (in the keV to few MeV range) scattering cross-sections of neutrons off various nuclei were experimentally studied by Fermi and his collaborators, and many other groups, around 1936. They found striking resonance structures in these cross-sections, with sharp maxima separated by very small values in between. Soon after, a theoretical explanation was offered independently by Niels Bohr on the one hand, and by Gregory Breit and Wigner[22] on the other. This is the so-called compound nucleus model. It pictures the scattering and reaction processes as taking place in two steps. At first the incoming low-energy projectile (which could be some light nucleus rather than a neutron) and the target combine to produce a compound nucleus in one of several possible metastable states. In this process the projectile energy is quickly shared with all the nucleons in the compound structure, and then the mode of formation of this structure is 'forgotten'. In the second step, the decay of the compound nuclear state into various energetically allowed channels is governed by probability laws. It is the probability of occurrence of the first step that shows an extremely sensitive energy dependence and gives rise to the observed resonances. In their work Breit and Wigner derived the famous bell-shaped single-level resonance formula known after their names:

Probability of formation of compound nucleus

$$\alpha\Gamma_\lambda / \{(E - E_\lambda)^2 + \tfrac{1}{4}\Gamma_\lambda^2\},$$

E = total initial energy,

$E_\lambda, \Gamma_\lambda$ = average energy, width, of compound nuclear state λ.

The partial cross-sections for subsequent decays into each of the several available final channels retain this characteristic energy dependence.

Sometime after this, around 1944, Fermi remarked to Wigner (as was mentioned earlier) that a good theoretical basis for the compound nucleus model was lacking. Thereupon Wigner set about formulating one. This was the starting point of the R-matrix theory of nuclear reactions, developed by him largely in collaboration with Eisenbud[23]. The basic idea is to separate the total multidimensional configuration space of all the nucleons in the compound nucleus (i.e. the projectile nucleons plus the target nucleons) into two parts: an interior region where they are *all* within the range of nuclear forces acting between every pair, and an exterior region where this is not so. In the latter region, one then defines or picks out essentially nonoverlapping subregions, one for each possible (two-body) final channel into which the compound nucleus can decay. Instead of posing a multichannel hamiltonian eigenfunction and eigenvalue problem, a series of matching conditions connecting the interior and exterior channel wave functions and their radial derivatives, across the borders between the interior and each exterior region, are set up. The R-matrix elements are quantities that enter these relations, they are a multichannel generalization of the logarithmic derivative of a wave function in a one-dimensional radial problem. The parameters entering the R-matrix are the energy values and the various partial decay widths of all possible compound nucleus levels. Thus, the R-matrix became simultaneously a convenient method for parametrization of scattering and reaction amplitudes using phenomenologically accessible compound nuclear state energies and widths, and with further developments, a way to embody general physical principles, such as unitarity and causality, governing reaction processes. *Inter alia* this led to a multilevel generalization of the Breit–Wigner resonance expression given above, and to a criterion for the validity of the single-level formula.

Returning to the problem of nuclear forces and structure, in 1937 Wigner came up with the SU(4) supermultiplet theory to systematize the low-lying energy levels of light nuclei[24]. The idea was that the interactions among protons and neutrons, regarded as nucleons possessing the isospin degree of freedom introduced by Heisenberg[25] as early as 1932, might to a good approximation be both spin- and isospin-independent. More generally, it might be invariant under all four-dimensional unitary transformations mixing up the four independent spin–isospin states of a nucleon. (This assumption actually leads to specific spin and isospin dependences in the interaction.) It would then be possible to arrange the energy levels of 'neighbouring' nuclei with a common mass number into various unitary irreducible representations (UIRs) of SU(4), consider systematically the breaking of this symmetry, etc. Each UIR of SU(4) is made up of several spin–isospin multiplets in a definite way. While the idea was physically well motivated as a useful first approximation, it was pursued only to a limited extent. However, many years later, in 1964, Wigner's theory provided the inspiration for a similar SU(6) invariant theory of baryons and mesons in the framework of the quark model [26].

At the other end of the scale from low-lying well-separated energy levels of light nuclei, we have the relatively highly excited and closely spaced levels of heavy nuclei with many degrees of freedom. Here Wigner proposed a completely different physical approach, one which has stimulated work by many others and led to connections with several other problems[27]. The physical ideas may be motivated as follows. As the excitation energy (of a complicated nucleus) increases, one expects the energy levels to get closer and closer, and one also loses hope of being able to obtain them individually from a first-principles Hamiltonian. Instead, what would be more accessible and physically interesting are various statistical properties of the levels; the probability distributions for successive levels to occur at various energies, for the spacing between neighbouring levels to have different values, and so on. To obtain these statistical features, and at the same time to reflect the fact that one is dealing with a very complex system with many degrees of freedom, Wigner proposed that the basic Hamiltonian (after truncation to a large but finite dimension) be itself regarded as a random matrix, belonging to an ensemble with specified properties. Once one specifies the nature of this ensemble, regarded as a primary input, the statistical properties of the eigenvalues of the Hamiltonian, the spacing distribution, etc., can all be derived, in principle, as secondary consequences. It turns out that in using this approach one must deal with one 'simple sequence' of nuclear levels at a time; this is a set of levels possessing the same exactly conserved quantum numbers – 'belonging to the same symmetry'– such as the total angular momentum and parity. Properties of different simple sequences are independent. Thus, Wigner's hypothesis was that the local statistical behaviour of the levels in a simple sequence is given by the properties of the eigenvalue spectrum of a random matrix drawn from a suitable ensemble. The type of ensemble to be used depends on the integer or

half odd integer nature of the total angular momentum, behaviour under time reversal, and presence or absence of rotational symmetry. Later work has shown that there are three natural types of ensembles, in correspondence with the three great families of classical compact simple Lie groups: the Gaussian real orthogonal, the Gaussian complex unitary, and the Gaussian symplectic ensembles. These ensembles consist respectively of real symmetric, complex hermitian and real quaternionic matrices (of suitable dimensions, even in the last case). The probability distribution defining the ensemble is invariant under a real orthogonal, complex unitary or unitary symplectic group of transformations applied to its elements; moreover, the matrix elements of the Hamiltonian are assumed to be independent random variables. It is the combination of these two properties that makes these ensembles Gaussian.

A great deal of sophisticated mathematical analysis has gone into these objects, and this activity continues[28]. One very interesting feature that was recognized very early was that the spacing distribution vanishes as a power of the spacing as the spacing tends to zero. The rate of this vanishing, the power involved, is characteristic for each of the three families of ensembles. The physical meaning of this result – borne out by experiments and reminding us of the no-crossing theorem of Wigner and von Neumann for electron terms of the same symmetry – is that within a simple sequence neighbouring levels do not like to come very close to one another. Had we imagined that the energy levels themselves were independently statistically distributed, there would have been no cause for such level repulsion. This only emphasizes Wigner's idea that the properties of the ensemble of Hamiltonians must be chosen first, and other properties then obtained as consequences.

Quantum field theory, relativistic classical and quantum mechanics

The rules for canonical quantization – creating a quantum theory from a classical one – were originally invented in the context of nonrelativistic particle quantum mechanics. The first successful application of these rules to a classical field theory came with Dirac's quantization of the electromagnetic field. This led to his classic 1927 paper in which he treated the processes of emission and absorption of radiation by matter, using quantum principles and the photon concept[29]. The quantized field led to a synthesis of complementary classical particle and field languages, and could describe states with variable numbers of identical particles. The canonical quantization method led to commutation relations of the form

$$a_r a_s^\dagger - a_s^\dagger a_r = \delta_{rs},$$

$$a_r a_s - a_s a_r = a_r^\dagger a_s^\dagger - a_s^\dagger a_r^\dagger = 0.$$

Here a_r ($a_r^\dagger$) are the annihilation (creation) operators for photons in various states indexed by r. These states are an independent, orthogonal and complete set of modes of the electromagnetic field. The operators a_r, $a_r^\dagger$ are quantum analogues of the classical complex coefficients in an expansion of the classical field in these modes. In this case the appearance of commutation relations led naturally to Bose–Einstein statistics for photons. Very soon after Dirac's paper, Jordan and Wigner showed that to describe fermions (such as electrons) obeying Pauli's exclusion principle and Fermi–Dirac statistics, the particle annihilation and creation operators must obey anticommutation relations[30]:

$$a_r a_s^\dagger + a_s^\dagger a_r = \delta_{rs},$$

$$a_r a_s + a_s a_r = a_r^\dagger a_s^\dagger + a_s^\dagger a_r^\dagger = 0.$$

For a finite number of modes, they proved that up to equivalence there is only one irreducible representation of these relations, and it is finite-dimensional. This uniqueness is similar to a corresponding result in the case of commutation relations. The major difference is that from a mathematical point of view systems of operators obeying the anticommutation relations are quite 'harmless', while in the case of commutation relations they are unbounded and the space is infinite-dimensional – even for a finite number of modes. Of course, in the Jordan–Wigner case there is no sensible classical limit.

It is interesting to note that Dirac's initial reaction to this work of Jordan and Wigner was decidedly negative[31]. Wigner later attributed this to Dirac's being very committed to the Hamiltonian point of view in dynamics – 'a captive of the Hamiltonian formalism'. However, it became clear very soon that this was the correct way to set up quantum field theory for fermions, and it became part of the foundations of the subject.

The first attempts at uniting quantum mechanics and special relativity were due to Klein and Gordon. This resulted in the wave equation named after them, but it faced problems of interpretation at the one-particle level. The next, spectacularly successful, attempt was Dirac's work in 1928 that led to his wave equation for the electron and its series of amazing consequences[32]. Probably soon after, in 1928 itself, Dirac suggested to Wigner a comprehensive study of all possible unitary irreducible representations of the inhomogeneous Lorentz group (IHLG), i.e. of the homogeneous Lorentz group (HLG) supplemented by space–time translations. By about 1932, Majorana had constructed many of these UIRs, and later these were simplified by Dirac and Proca[33]. The solution of this problem posed by Dirac to Wigner became a herculean effort, being completed only in 1937. The result was an all-time classic paper in mathematical physics[34]. In it, Wigner acknowledges the help and guidance he received not only from Dirac but also on mathematical aspects from von Neumann. At some stage Dirac advised Wigner to be careful, and the latter replied[35]: 'You point out that care is needed in the analysis of the representations of the Lorentz group; I promise you that I will be careful'.

Wigner's paper contains a detailed analysis of the structure of the HLG and the IHLG, and of general unitary representations (URs) of the IHLG in the context of quantum mechanics; it then turns to a study of the UIRs. The result was that these could be classified into four broad types, depending upon the nature of the possible values of energy–momentum p^μ occurring within the UIR, and the allowed 'states of polarization' for each energy–momentum. The helicity λ is defined as the component of angular momentum in the direction of momentum. For each kind of p^μ (provided it is not identically vanishing) the allowed values of λ are determined by some UIR of a corresponding subgroup of the HLG, the so-called 'little-group' for that p^μ; it consists of all elements of the HLG which leave p^μ invariant. The pattern of UIRs

Table 1. The pattern of UIRs of the IHLG

Nature of p^μ	Little group within HLG(SL(2,C))	Number of polarization states, spectrum of λ	Remarks
(a) Time-like (positive or negative)	SO(3)(SU(2))	$2s+1$ for $s = 0, 1/2, 1 \ldots$ $\lambda = s, s-1, \ldots, -s$	Massive particles with zero or finite spin $s = 0$ for π meson $s = 1/2$ for electron
(b) Light-like positive or negative)	E(2), two-dimensional Euclidean group	One: $\lambda = 0$ Two: $\lambda = \pm s$, $s = 1/2, 1, \ldots$ Infinite: $\lambda = 0, \pm 1 \pm 2, \ldots$ or $\lambda = \pm 1/2, \pm 3/2, \ldots$	No known particles $s = 1$ for photons $s = 1/2$, $\lambda = -1/2$ for neutrinos No known particles
(c) Space-like	SO(2,1) (SL(2,R))	One: $\lambda = 0$ Infinite: $\lambda = s, s+1, \ldots$ or $-s, -s-1, \ldots$ $s = 1/2, 1, \ldots$ or $\lambda = 0, \pm 1, \pm 2, \ldots$ or $\lambda = \pm 1/2, \pm 3/2, \ldots$	Imaginary mass, unphysical
(d) Vanishing	HLG(SL(2, C))	–	–

of the IHLG is displayed in Table 1 (here space inversion or parity has been included in the HLG, except that for neutrinos this operation is undefined)[36].

While many of these UIRs were known earlier to Majorana and Dirac, the so-called infinite-spin or continuous-spin representations in cases (b) and (c) were genuinely new. In his work, Wigner did not carry the investigation of these, or of case (d), to completion. He mentioned their existence, and only remarked: '... the last case may be the most interesting from the mathematical point of view. I hope to return to it in another paper. I did not succeed so far in giving a complete discussion of the 3rd class.' Wigner's 'last case' and '3rd class' correspond respectively to (c) and (d) in our table. We also see that not every mathematically acceptable UIR of the IHLG is acceptable on physical grounds.

Relativistic quantum systems described by any UIR of the IHLG are called 'elementary systems'. Truly elementary particles, able to exist in isolation, are described using them. Examples are photons, neutrinos, electrons and muons. The phrase 'elementary systems' conveys the meaning that all their properties are revealed by studying their behaviour under all elements of the IHLG – there is no internal structure involved. In the above listing, only cases (a) and (b) for finite helicity are realized in nature.

The UIRs of case (d) are actually UIRs of the HLG SO (3,1) (or of the closely related group SL *(2, C)*). It remained for Harish-Chandra and for Gel'fand and Naimark to determine them independently[37]. The inputs needed to construct the UIRs of case (c) for infinite spin were provided by Bargmann through his construction of the UIRs of SO (2,1) and SL $(2,R)$[38].

In his contribution to RMP, Dirac made the following comments on Wigner's work[39]. 'The problem of working out all unitary representations of the IHLG has been dealt with by Wigner, taking the mathematical point of view that two representations are equivalent if they are connected by a unitary transformation. He decomposes the representations into their irreducible constituents and finds that the irreducible constituents provide theories of elementary particles with various spins. This work does not lead to any interaction between particles. To bring in interaction, one must depart from the point of view of looking at two representations as equivalent if they are connected by a unitary transformation, a point of view which involves looking upon all unitary transformations as trivial. To a physicist, some unitary transformations are trivial, whereas others (for example, the S matrix) are far from trivial, so a physicist cannot look upon two representations connected by a unitary transformation as necessarily equivalent.' The point is that for any really interesting relativistic quantum system, such as a relativistic quantum field theory, it is not only important to know which UIRs of the IHLG are present, but also how they are put together. However it must be pointed out that as early as 1949 Wigner himself had drawn attention to this situation[40]: 'The elementary systems correspond mathematically to irreducible representations of the Lorentz group and as such can be enumerated... However, in the description by irreducible states, the form of almost all physically important operators remains unknown and, in fact, depends on the system, the types of interactions, etc. This leads to a rather strange dilemma: in the customary description the form of the physically important operators is known but the time dependence of the states is unpredictable or difficult to calculate. In the description just mentioned, the situation is opposite: the time dependence of the

states follows from the invariance properties, but the form of the physically important operators is hard to establish.'

Wigner returned on many occasions to a description of the results of his classic work. He also constructed with Bargmann a unified set of wave equations whose solutions would lead to UIRs of types (a) and (b) in our table[41]. His work with Newton on the problem of position is particularly interesting, so I describe it in a little detail [42].

The starting point of nonrelativistic particle quantum mechanics is the set of positions and momenta as primary dynamical variables, out of which all other variables are built up. (Later work has shown that these positions and momenta can be derived as secondary objects starting from suitable quantum-mechanical representations of the Galilei group). Now, from Wigner's point of view in the relativistic context, the primary things are the UIRs of the IHLG. After having set them up, one must examine within which UIRs one can construct position operators with physically desirable properties. Such an analysis was first undertaken by Newton and Wigner. They were able to show that in every finite mass and finite spin UIR (case (a)) a unique set of position operators possessing several physically reasonable properties does indeed exist. However, contrary to naive expectation, they do not form the space components of a relativistic four-vector. This illustrates the fact that in quantum theory the unitary transformation law is more basic than the geometric one or manifest covariance. In the massless case with finite nonzero helicity even this much cannot be done. Thus, neither photons nor neutrinos can be localized in space.

In other related work we mention the study by Inonu and Wigner of the process of 'group contraction' by which the IHLG goes over in the nonrelativistic limit to the Galilei group[43]; Salecker and Wigner's analysis of deep conceptual problems in bringing together quantum mechanics and general relativity, caused by quantum limitations on position measurements[44]; and van Dam and Wigner's construction of classical relativistic direct-interaction theories resting upon integrodifferential equations for particle trajectories[45]. One of Wigner's conclusions was that while special relativity and quantum mechanics could at least conceptually be combined, with general relativity and quantum mechanics there was no common ground at all.

Interpretation of quantum mechanics

In the early 1960s Wigner turned to a serious examination of the problems of interpretation of quantum mechanics, and a clear expression of the orthodox position which essentially coincided with his own[46]. As evidence for the latter, here is his own statement: 'The orthodox view is very specific in its epistemological implications ... A large group of physicists finds it difficult to accept these conclusions and, even though this does not apply to the present writer, he admits that the far-reaching nature of the epistemological conclusions makes one uneasy.' He also often said that he was adding hardly anything new to London and Bauer's classic 1939 exposition[47]. He accepted the treatment of measurement theory that had been articulated by his friend von Neumann[48] as early as 1932, and wanted to restate it for a new generation and extract its ultimate consequences for epistemology.

Wigner emphasized that the state vector of a quantum system changes in two mutually exclusive ways – continuous, deterministic Schrödinger evolution when not subject to observation, and discontinuous, probabilistic, collapse when measurements are made. He went to much length to show that the linear Schrödinger equation – even including the apparatus and the system's coupling to it – can never produce the macroscopically desired collapse phenomenon, and stressed repeatedly that pure states and mixtures have very different physical properties. He also presented a pragmatic answer to the question 'What is the state vector?'. It was that it codifies in a compact way all past information about a system, on the basis of which we can state the probabilistic connections that quantum mechanics gives among a series of measurements carried out subsequently and sequentially in time; all the consequences of quantum mechanics are just such statements. So, as the orthodox view claims, 'the laws of quantum mechanics can be expressed only in terms of probability connections', and cannot be formulated in terms of objective reality.

Pursuing this analysis further, Wigner came to the conclusion that human consciousness is an essential external ingredient needed to make complete sense of quantum mechanics. The collapse of the state vector occurs when and only when an observation is registered in some individual consciousness: 'It is at this point that the consciousness enters the theory unavoidably and unalterably. If one speaks in terms of the wave function, its changes are coupled with the entering of impressions into our consciousness'. And again: '... it was not possible to formulate the laws of quantum mechanics in a fully consistent way without reference to the consciousness'. In support of this declaration, Wigner appeals to Heisenberg and says: 'W. Heisenberg expressed this most poignantly (*Daedalus*, 1958, **87**, 99): "The laws of nature which we formulate mathematically in quantum theory deal no longer with the particles themselves but with our knowledge of the elementary particles ... The conception of objective reality... evaporated into the... mathematics that represents no longer the behaviour of elementary particles but rather our knowledge of this behaviour" '.

As one can imagine, this line of thinking led Wigner inexorably to a kind of solipsism, and to the delineation of two kinds of reality – the content of one's own consciousness, the only absolutely real, and everything else external to oneself, including every other person's consciousness. To support the former he turned to Schrödinger; '... the most eloquent statement of the prime nature of the consciousness with which this writer is familiar and which is of recent date is on page 2 of Schrödinger's *Mind and Matter*: "Would it (the world) otherwise (without consciousness) have remained a play before empty benches, not existing for anybody, thus quite properly not existing?" ' But there was a sign of asymmetry – the only absolutely real, one's own consciousness, does depend on food, air and water for its own survival and functioning, as we are painfully aware; so he made a case for devising experiments which might show up the effects of consciousness on matter. In talking of the first kind of reality, Wigner also realized and stated its obvious limitations – its awakening with birth and infant growth, its extinction at death. So he argued for a deep study of

the former phase, to understand the nature of consciousness.

Wigner felt that the development of quantum mechanics had widened the outlook of most physicists, and also in a sense made them inward-looking: 'Until not many years ago, the "existence" of a mind or soul would have been passionately denied by most physical scientists ... Even today, there are adherents to this view though fewer among the physicists than – ironically enough – among biochemists'. He also saw that quantum mechanics reinforces the circumstance that any observation and interpretation of measurement rests on previously constructed and understood theory. Thus, we are linked in a chain to the very beginnings of our acquisition of knowledge of our surroundings and its regularities – indeed to phylogenesis and ontogenesis.

Today it may seem that these conclusions of Wigner were premature. Certainly, efforts are aplenty to find more 'acceptable' interpretations of quantum mechanics, without appeal to ourselves as essential prerequisites. Was Wigner then 'a victim of his generation'? Should we smile at these conclusions which he found inescapable? Or was he only being ruthlessly honest and expressing clearly what others hesitated to put into words?

Solid-state physics, reactor theory and technology

I will touch upon these areas only briefly. Wigner's interest in problems of solid-state physics and materials science stemmed from a very early date. There must have been links to his original training as a chemical engineer; later on his detailed knowledge of properties of materials played a key role in his work on reactors. Among his gifted students in solid-state science in the 1930s we may mention John Bardeen, Gregory Wannier and Frederick Seitz. It was Wigner who suggested to Wannier[49] 'that there ought to be a way to reconcile the local and the band concept for electrons, and that such a reconciliation would probably be useful in understanding the spectra of insulators'. Wigner also worked on radiation damage in solids – the detailed microscopic picture of lattice defects occurring when materials are irradiated with neutrons, the resulting changes in heat and electrical conductivity and ductility, and also the ways in which the material seems to recover from the damage as time goes on[50].

Wigner was the source of much of the theory and the major technological developments connected with nuclear reactors. His contributions began in 1940. As briefly mentioned earlier, he was a leader at the University of Chicago Metallurgical Laboratory during 1943–45. He contributed to the development of research reactors, power reactors and plutonium production reactors. On the theoretical front he made major contributions to the spectrum of the Boltzmann equation, neutron thermalization, thermal utilization and resonance absorption. All very practical contributions 'which one would hardly, *a priori*, have associated with the same man who introduced group theory into quantum mechanics'[51].

Views on science, philosophy and life

Wigner was a gifted and articulate expositor of science and its principles to general audiences. However, he frequently indulged in a kind of mock humility – as his Princeton colleagues explained his language[52], 'A piece of work is "amusing" if it is correct and beautiful; it is "interesting" if it is wrong and messy.' And in describing the epistemology of quantum mechanics to an audience of nonphysicists, he said of himself the writer[53]: 'He realizes the profundity of his ignorance of the thinking of some of the greatest philosophers and is under no illusion that the views to be presented will be very novel. His hope is that they will appear sensible.' He could convey sharp ideas pithily: 'Some one once said that philosophy is the misuse of a terminology which was invented just for this purpose'.

These apart, his grasp of and concern for the grand principles of science were very deep. The role of invariance principles and their associated conservation laws captivated him – he dwelt upon them at length on many occasions[54], and said: 'A large part of my scientific work has been devoted to the study of symmetry principles in physics ....' He titled his Nobel lecture 'Events, laws of nature, and invariance principles'. He often described as a miracle the fact that human understanding could uncover laws of nature, and separate them from the accidents of initial conditions. The laws provide structure and coherence to events, and, in turn, the symmetry principles provide these qualities to laws; thus one has the ascending progression: events to laws to symmetry principles.

Turning to the role of mathematics in natural science, he expressed wonder at the way in which mathematical concepts and connections show up in unexpected ways and places, and also at the fact that tentative theories turn out upon further development to be far more accurate than could reasonably have been expected at the outset. This led him to conclude that, since we do not quite know why we succeed so well so often, we must be cautious and not immediately regard a successful explanation as the truth!

Pondering on the likely future of science, Wigner wondered whether it might not wind down under its own weight, and lose its attractiveness to the young. The increasing extent of science makes it go beyond the reach of any one individual. But the response to this cannot just be an increase in team efforts, because this can never capture true creative thinking in the individual subconscious. There is a need here to find deeper ways of sharing information and insight, of harmonizing the collective conscious with the subconscious in each individual.

Continuing on the theme of the growth of science and the emergence of large collaborative efforts, he argued for protecting the individual and giving value and esteem to little science: 'One does not have the satisfaction which creative work, as we know it today, provides, if one's activities are too closely directed by others'. About the emergence of deep insights, 'It is hard to imagine how they can be developed other than in comparative solitude'. And as for the pleasures of pursuing science: 'It has been said that the only occupations which bring true joy and satisfaction are those of poets, artists, and scientists, and, of these, the scientists are apparently the happiest.'

Through the description of his work I have tried to convey the fact that Wigner acknowledged very graciously his debt to some of his most gifted contemporaries. He was also generous in his assessment of them. Of von Neumann he wrote: '... whenever I talked with the sharpest intellect whom I have known – with von Neumann – I always had the impression that only he was fully awake, that I was halfway in a dream.' And about Richard Feynman:

'He is a second Dirac, only this time more human'.

Two persons that Wigner had been very close to – Enrico Fermi and von Neumann – both died in their fifties. Wigner described and contrasted their attitudes to the inevitable. With Fermi, 'On a heroic scale was his acceptance of death ... He was so completely composed that it appeared superhuman'. But with von Neumann it was very different: 'It was heartbreaking to watch the frustration of his mind, when all hope was gone, in its struggle with the fate which appeared to him unavoidable but unacceptable'. These experiences must have affected Wigner deeply; at a convocation address to an audience of young students soon after, he said: 'Our culture is committing a sin by covering our eyes against the realization that none of us will be here always'. And to a general audience some time later: 'The recognition that physical objects and spiritual values have a very similar kind of reality has contributed in some measure to my mental peace'. These various expressions seem related.

Wigner was a physicist who achieved rare range and depth in his life and work. He was a product of the old world who flowered during the golden age of theoretical physics, and carried the fragrance of his subject to the new world. The time seems past when such another can appear.

References and footnotes

1. In the following, frequent references will be made to the following three sources: (a) Wigner, E. P., *Group Theory and its Application to the Quantum Mechanics of Atomic Spectra*, Academic Press, New York, 1959; (b) *Reviews of Modern Physics*, vol. 34, no. 4, (c) Wigner E. P., *Symmetries and Reflections – Scientific Essays*, Indiana University Press, 1967. For brevity we shall refer to these as GT, RMP and SR, respectively.
2. SR, p. 257.
3. Wigner, E. P., Remembering Paul Dirac, in *Reminiscences about a Great Physicist: Paul Adrien Maurice Dirac*, (eds. Kursunoglu, B. N. and. Wigner, E. P.), Cambridge University Press, Cambridge, 1987.
4. Dirac Margit, Thinking of my darling Paul, in *Reminiscences about a Great Physicist: Paul Adrien Maurice Dirac*, (eds. Kursunoglu, B. N. and. Wigner E. P., Cambridge University Press, Cambridge, 1987.
5. Wheeler, J. A., in RMP, p. 873.
6. Dresner, L. and Weinberg, A. M., in RMP, p. 747.
7. Wigner, E. P., Über die Erhaltungssatze in der Quanten mechanik, *Nachr. Ges. Wiss. Göttingen, Math – Physik*, K1. 1927, p. 375, SR, p. 61.
8. Laporte, O., *Z. Phys.*, 1924; **23**, 135; GT, Chap. 18.
9. GT., Chaps. 15 and 16; Bargmann, V., in RMP, p. 829; Biedenharn, L. C. and Louck, J. D., *Angular Momentum in Quantum Physics – Theory and Application, Encyclopedia of Mathematics and its Applications* (ed. Rota, G. C.), Addison-Wesley, Reading MA, 1981, vol. 8.
10. GT, Appendix to Chap. 20; Bargmann, V., *J. Math. Phys.*, 1964, **5**, 862; Biedenharn, L. C. and Louck, J. D., *The Racah-Wigner Algebra in Quantum Theory, Encyclopedia of Mathematics and its Applications*, (ed. Rota, G. C.), Addison-Wesley, Reading, MA, 1981, vol. 9, Chap. 5.
11. Wigner, E. P., Über die Operation der Zeitumkehr in der Quanten mechanik, *Nachr. Ges. Wiss. Gottingen, Math-Physik*, K1., 1932, p. 546; GT, Chap. 26.
12. von Neumann, J. and Wigner, E. P., Über das Verhalten von Eigenwerten bei adiabatischen Prozessen, *Phys. Z.*, 1929, **30**, 467; see also Landau, L. D. and Lifshitz, E. M., *Quantum Mechanics*, Pergamon Press, Oxford, 1965, p. 279.
13. Wick, G. C., Wightman, A. S. and Wigner, E. P., The intrinsic parity of elementary particles, *Phys. Rev.*, 1952, **88**, 101; see also Streater, R. F. and Wightman, A. S., *PCT, Spin and Statistics, and all that*, Benjamin, W. A., New York, 1964, Chap. 1.
14. GT, p. v.
15. Weisskopf, V. and Wigner, E. P., Berechnung der naturlichen Linienbreite auf Grund der Diracschen Licht theorie, *Z. Phys.*, 1930, **63**, 54; Über die naturliche Linien breite in der Strahlung des harmonischen Oszillators, *ibid.*, 1930, **65**, 18; see also Goldberger, M. L. and Watson, K. M., *Collision Theory*, Wiley, New York, 1964, Chap. 8; Merzbacher, E., *Quantum Mechanics*, Wiley, New York, 1970, 2nd edn., chap. 18.
16. Wigner, E. P., On the quantum correction for thermodynamic equilibrium, *Phys. Rev.*, 1932, **40**, 749; see also Hillery, M., O'Connell, R. F., Scully, M. O. and Wigner, E. P., Distribution functions in physics: Fundamentals, *Phys. Rep.*, 1984, **106**, 121.
17. Weyl, H., *The Theory of Groups and Quantum Mechanics*, Dover, New York, 1931, p. 274; Moyal, J. E., *Proc. Camb. Phil. Soc.*, 1949, **45**, 99.
18. Wigner, E. P., Lower limit for the energy derivative of the scattering phase shift, *Phys. Rev.* 1955, **98**, 145; see also Goldberger, M. L. and Watson, K. M., in ref. 15, chap. 8, p. 492; Brenig, W. and Haag, R., General quantum theory of collision processes, in *Quantum Scattering Theory* (ed. Mare Ross.), Indiana University Press, Bloomington, 1963, Section 3.
19. von Neumann, J. and Wigner, E. P., Über merkwurdige diskrete Eigenworte, *Phys. Z.*, 1929, **30**, 465; see also Ballentine, L. E., *Quantum Mechanics*, Prentice-Hall, Englewood Cliffs, NJ, 1990, p. 205.
20. Jordan, P., von Neumann, J. and Wigner, E. P., On an algebraic generalization of the quantum mechanic formalism, *Ann. Math.*, 1934, **35**, 29; see also Biedenharn, L. C. and Louck, J. D., in ref. 10, chap. 5.
21. Wigner, E. P., On the mass defect of helium, *Phys. Rev.*, 1933, **43**, 252; Über die streuung von neutronen an protonen, *Z. Phys.*, 1933, **83**, 253; Bartlett Jr., J. H., *Phys. Rev.*, 1936, **49**, 102; Majorana E., *Z. Phys.*, 1933, **82**, 137; Heisenberg, W., *Z. Phys.*, 1932, **77**, 1.
22. Bohr, N., *Nature* 1936, **137**, 344; Breit, G. and Wigner, E. P., Capture of slow neutrons, *Phys. Rev.*, 1936, **49**, 519.
23. Wigner, E. P., Resonance reactions and anomalous scattering, *Phys. Rev.*, 1946, **70**, 15; Resonance reactions, *Phys. Rev.* 1946, **70**, 606; Eisenbud, L. and Wigner, E. P., Higher angular momenta and low-range interaction in resonance reactions, *Phys. Rev.*, 1947, **72**, 29; SR, p. 93; Vogt, E. in RMP, p. 723.
24. Wigner, E. P., On the consequences of the symmetry of the nuclear Hamiltonian on the spectroscopy of nuclei, *Phys. Rev.*, 1937, **51**, 106; reprinted in Dyson, F. J., *Symmetry Groups in Nuclear and Particle Physics*, Benjamin, W. A., New York, 1966.
25. Heisenberg, W., *Z. Phys.*, 1932, 77, 1.
26. See, for instance, the reprint group 3 in Dyson, F. J., ref. 24.
27. Wigner, E. P., Gatlinberg Conf. on Neutron Physics, Oak Ridge Natl. Lab. Rept. No.ORNL-2309, p. 59; On the statistical distribution of the widths and spacings of nuclear resonance levels, *Proc. Camb. Phil. Soc.*, 1951, **47**, 790.
28. Dyson, F. J., *J. Math. Phys.*, 1962, **3**, 140, 1191; Dyson F. J. and Mehta, M. L., *J. Math. Phys.*, 1963, **4**, 701; Mehta, M. L., *Random Matrices*, Academic Press, New York, 1967.
29. Dirac, P. A. M., *Proc. R. Soc. London*, 1927, **A114**, 243.
30. Jordan, P. and Wigner, E., Über das Paulische Aquivalenzverbot, *Z. Phys.*, 1928, **47**, 631.
31. Kragh, H. S., *Dirac – A Scientific Biography*, Cambridge University Press, Cambridge, 1990, pp. 128–129, 289, 338.
32. Dirac, P. A. M., *Proc. R. Soc. London*, 1928, **A117**, 610; **A118**, 351.
33. Majorana, E., *Nuovo Cim.*, 1932, **9**, 335; Dirac P. A. M., *Proc. Roy. Soc.*

London, 1936, **A155**, 447; Al. Proca, *J. Phys. Rad.*, 1936, 7, 347.
34. Wigner, E., On unitary representations of the inhomogeneous Lorentz group, *Ann. Math.*, 1939, **40**, 149.
35. Quoted in ref. 13.
36. For quantum-mechanical purposes, one has to deal with the group $SL(2, C)$, the universal covering group of the HLG, and its subgroups.
37. Harish-Chandra, *Proc. Roy. Soc. London*, 1947, **A189**, 372; Gel'fand I. M. and Naimark, M. A., *Izv. Akad. Nauk. SSSR*, 1947, **11**, 411.
38. Bargmann, V., *Ann. Math.*, 1947, **48**, 568.
39. Dirac, P. A. M., *in* RMP, p. 592.
40. Wigner, E. P., Invariance in physical theory, *Proc. Am. Phil. Soc.*, 1949, **93**, 521; reprinted in SR, p. 8-9.
41. Bargmann, V. and Wigner, E. P., Group theoretic discussion of relativistic wave equations, *Proc. Natl. Acad. Sci., USA*, 1948, **34**, 211.
42. Newton, T. D. and Wigner, E. P., Localized states for elementary systems, *Rev. Mod. Phys.*, 1949, **21**, 400.
43. Inonu, E. and Wigner, E. P., On the contraction of groups and their representations, *Proc. Natl. Acad. Sci. USA*, 1953, **39**, 510.
44. Salecker, H. and Wigner, E. P., Quantum limitations of the measurement of space–time distances, *Phys. Rev.*, 1958, **109**, 571; SR, pp. 62 ff.
45. Wigner, E. P. and van Dam, H., Classical relativistic mechanics of interacting point particles, *Phys. Rev.*, 1965, **B138**, 1576.
46. Wigner, E. P., The problem of measurement, *Am. J. Phys.*, 1963, **31**, 6; Remarks on the mind–body question, in *The Scientist Speculates* (ed. Good, I. J.), William Heinemann, London, 1961; Two kinds of reality, *The Monist*, 1964, **48**, (2); all reprinted in SR, pp. 153, 171, 185.
47. London, F. and Bauer, E., *La Theorie de l'Observation en Mecanique Quantique*", Hermann et Cie, Paris, 1939.
48. von Neumann, J., *Mathematical Foundations of Quantum Mechanics*, Princeton University Press, Princeton, New Jersey, 1955.
49. Wannier, G. in RMP, p. 645.
50. SR, p. 126.
51. Dresner Lawrence and Weinberg, A. M., in RMP, p. 747.
52. Bargmann, V. *et al.*, in RMP, p. 588.
53. SR, p. 186.
54. In addition to the articles included in SR, see also Houtappel, R. M. F., van Dam, H. and Wigner, E. P., The conceptual basis and use of the geometric invariance principles, *Rev. Mod. Phys.*, 1965, **37**, 595.

N. Mukunda is in the Centre for Theoretical Studies and Department of Physics, Indian Institute of Science, Bangalore 560 012, India.

INDIAN INSTITUTE OF SCIENCE
Bangalore 560 012

Postdoctoral Research Associateships in Life Sciences and Biotechnology
(A DBT-sponsored programme)

Applications are invited from candidates with PhD in biological, chemical or physical sciences, or MD, for the award of temporary Research Associateship (maximum duration 2 years) to carry out research in frontier areas of biology in the Division of Biological Sciences at IISc under the DBT-sponsored postdoctoral programme. Candidates who have submitted their theses are also eligible to apply, but they may be appointed as Provisional Research Associates till they obtain the degree. Minimum consolidated emoluments for Research Associates will be around Rs 4,000 p.m.

Candidates should apply on plain paper, giving (1) biodata, (2) a summary of the PhD work, including the list of publications (reprints if any), (3) two reference letters, from the supervisor and from persons familiar with their scientific work, and (4) an undertaking that they will continue with the training programme for its duration.

After initial screening, candidates will be called for an interview at IISc, Bangalore, in the third/fourth week of September 1995. The PDF programme will start from 1 October 1995. Send applications to Prof. K. P. Gopinathan, Convenor, DBT-PDF programme at IISc Chairman, Department of Microbiology and Cell Biology, Indian Institute of Science, Bangalore 560 002, to reach him before **10 September 1995**.

Edited and published by Prof. S. Ramaseshan and Prof. P. Balaram, Current Science Association, Bangalore 560 080.
Typeset by Creative Literati Pvt. Ltd., Bangalore (Tel. Cum Fax: 5462698) Printed at Printek, Bangalore (3357763)

We'll help you get there

If you're selling science books/journals or laboratory products, or wish to recruit scientists in your establishment, then

you're sure to benefit by advertising here.

Because *Current Science* reaches nearly every university, scientific institution and industrial R&D unit in India. What's more, it's read by hundreds of individual subscribers – students, doctoral scientists and professionals in virtually every field of scientific activity in India.

Current Science has the largest circulation among scientific journals in India.

And to give you more impact, we're now bigger, and better.

Write now, or send your copy to

CURRENT SCIENCE

C.V. Raman Avenue, P.B. No. 8001,
Bangalore 560 080.

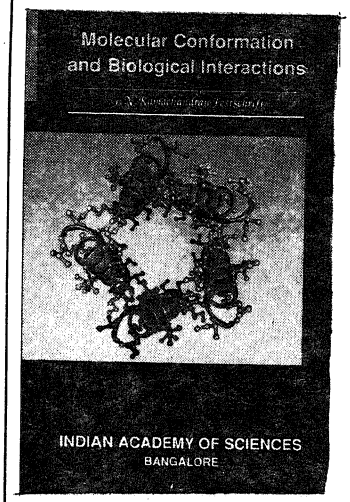
NEW Display Advertisement Rates*

		Tariff (rupees)					
		Inside pages		Inside cover pages		Back cover page	
No. of insertions	Size	B&W	Colour	B&W	Colour	B&W	Colour
1	Full page	5,000	10,000	7,000	12,000	10,000	15,000
	Half page	3,000	5,000	–	–	–	–
6	Full page	25,000	50,000	35,000	60,000	50,000	75,000
	Half page	15,000	25,000	–	–	–	–
12	Full page	50,000	1,00,000	70,000	1,20,000	1,00,000	1,50,000
	Half page	30,000	50,000	–	–	–	–
24	Full page	90,000	1,80,000	–	–	–	–
	Half page	50,000	90,000	–	–	–	–

SEND PAYMENT BY BANK DRAFT WITH ORDER

*Effective from 1 January 1995

3 SCHOLARLY PUBLICATIONS
from the INDIAN ACADEMY OF SCIENCES

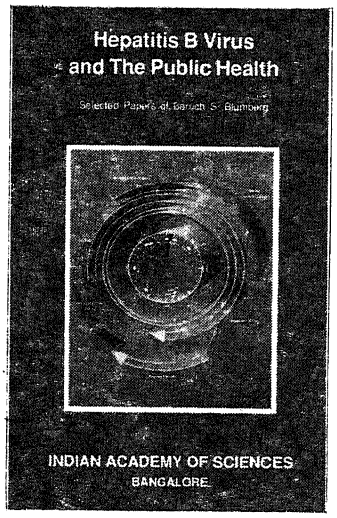


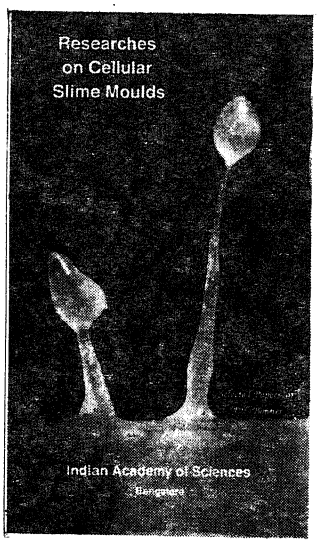
Molecular Conformation and Biological Interactions
G. N. Ramachandran Festschrift

This tribute to G. N. Ramachandran, one of the pioneers in the area of protein structure and conformational analysis, brings together over 40 state-of-the-art articles on protein structure and function, modelling and design of novel proteins, membrane channels, drug design and molecular recognition, nucleic acid structures.

Hepatitis B Virus and the Public Health
Selected papers of Baruch S. Blumberg

Blumberg received the Nobel prize for physiology or medicine in 1976 for his researches on hepatitis B virus and his discoveries of mechanisms relating to the origin and dissemination of infectious diseases. This collection of papers covers a wide range of fascinating topics—virus subtypes, sex ratio, insect vectors, cancer, epidemiology, polymorphisms.





Researches on Cellular Slime Moulds
Selected papers of John Tyler Bonner

Dictyostelium discoideum, the best-studied species of the cellular slime moulds, or social amoebae, is a favourite organism for work dealing with a range of phenomena studied by molecular, cellular and developmental biologists. Perhaps no single person has enriched our understanding of the cellular slime moulds as much as J. T. Bonner.

CIRCULATION DEPARTMENT
INDIAN ACADEMY OF SCIENCES
C. V. RAMAN AVENUE
P. B. NO. 8005
BANGALORE 560 080, INDIA

INFORMATION FOR CONTRIBUTORS

GENERAL

All manuscripts should be addressed to the Editor, *Current Science*, P. B. No. 8001, C. V. Raman Avenue, Bangalore 560 080. Submission of an article will be held to imply that it has not been previously published and is not under consideration for publication elsewhere; and further, that if accepted, it will not be published elsewhere. *Three copies of contributions of all categories* are required, with a letter of transmittal giving (i) names and complete addresses of the authors and (ii) title of the contribution and the category in which it is submitted (see below).

Current Science is a multidisciplinary journal and therefore research and review papers of general significance that are written clearly and well organized will be given preference. All papers will be first assessed by a Reviewing Editor. Papers found unsuitable in terms of the overall requirements of the journal will be returned to the authors. The others will be sent for detailed review. *Both solicited and unsolicited material will be reviewed.* Authors of these papers will be notified of acceptance, rejection, or need for revision of the paper. Returned papers cannot be resubmitted.

Illustrations and other material to be reproduced from other publications must be properly credited; it is the authors' responsibility to obtain permission for reproduction (copies of letters of permission should be sent).

CATEGORIES OF MANUSCRIPT

General articles (not exceeding 5000 words) discuss current trends in research in a field that will be of interest to readers outside the field; interdisciplinary topics: science policy and administration; or some aspect of the application of science and technology to human needs or the impact of science and technology on society/ecosystems/life. They should include a summary not exceeding 100 words, introductory paragraph(s), brief subheads at appropriate places to point to what follows, illustrations that will help a general reader, and references.

Review articles (not exceeding 5000 words) are expected to survey and discuss recent developments in a field. They should be well focused and organized, and avoid a general, 'textbook' style.

Research articles (not exceeding 4000 words) should report research results of fairly major significance. They should include an abstract not exceeding 100 words, introductory paragraph(s), and brief subheads.

Research communications (not exceeding 2000 words) should contain important findings that are novel and of fairly broad interest. They should include a brief abstract and an introductory paragraph. Text should not be broken up under subheads.

Correspondence includes letters, not exceeding 500 words, that are of general interest to scientists. All letters cannot be published.

Scientific correspondence contains technical comments, including those on articles or communications published in *Current Science* within the previous six months. Letters may be reviewed and edited.

Research news articles are intended to inform nonspecialists about recently published advances or important findings discussed at a meeting. Authors should also send a copy of the paper(s) on which the article is based. Meeting reports should avoid merely listing brief accounts of topics discussed, and must convey to readers the significance of an important advance.

Research accounts articles are intended to be personalized reviews of research from the authors' own laboratory, based on a body of published work. The articles must provide appropriate background to the area in a concise introduction, which should also serve to place the author's work in proper perspective. Articles will normally not exceed 8 to 10 printed pages.

Opinion articles present views on issues related to science and scientific activity. **Commentary** articles should contain expository notes on issues related to science and scientific activity.

Book reviews. Unsolicited reviews will also be considered. Reviews that merely 'list' brief descriptions of the contents cannot be published. Reviews should have 'context' and convey some information about the subject of the book.

Historical commentary and notes inform readers about interesting aspects of personalities or institutions of science or about watershed events in the history/development of science; most are expected to relate to India. Illustrations are welcome. Brief items will also be considered.

MANUSCRIPT PREPARATION

Manuscripts should be typed double-spaced on one side of white bond paper (21 × 28 cm). The pages should be numbered consecutively, starting with the title page and through the text, reference list, tables and figure legends. The **title** should be brief, specific and amenable to indexing. Not more than five **keywords** should be indicated separately; these should be chosen carefully and must not be phrases of several words. **Summary** and **abstract** should not have more than 100 words and should convey the main point of the paper, outline the results and conclusions, and explain the significance of the results.

Text. All papers should have a brief introduction. The text should be intelligible to readers in different disciplines and technical terms should be defined. Tables and figures should be referred to in numerical order. All **symbols** and **abbreviations** must be defined, and used only when absolutely necessary. Superscripts and subscripts and ambiguous characters should be clearly indicated. **Units of measure** should be metric or, preferably, SI. Methods should, as far as possible, be described briefly in appropriate table and figure legends.

Figures. In the case of line drawings one set of originals (without any lettering) is sufficient, with two copies containing lettering. In the case of photographs good prints are required with each copy of the manuscript; photocopies are not acceptable. Line drawings should be roughly twice the final printed size. The correct orientation should be indicated if not clear.

Photomicrographs and other photographs that require it must have a scale bar, which should be defined clearly in the legend. Primary data should be submitted as far as possible (e.g. actual photographs of electrophoretic gels rather than idealized diagrams).

References should be numbered in the order in which they appear, first through the text and then through table and figure legends. The following are examples of ways of writing references.

1. Mukundan, T. and Kishore, K., *Curr. Sci.*, 1991, **60**, 355–362.
2. Constantine, G., in *Biology of Bats* (ed. Wimsatt, W. A.), Academic Press, New York, 1970, vol. 1, pp. 319–322.

Acknowledgements should be brief. Footnotes are not allowed except to identify the corresponding author if not the first.

Cover photographs. Good photographs (colour or black and white) that pertain to a submitted paper will be considered for use on the cover. Good *prints* and a legend should be submitted with the manuscript. In the case of a colour picture, a transparency will be required for printing if accepted.

PROOFS AND PUBLICATION

Two sets of galley proofs are sent to the corresponding author. A reprint order form accompanies the proofs.

CURRENT SCIENCE

A fortnightly journal of research

Editors: S. Ramaseshan and P. Balaram

Editorial Board

S. Arunachalam, Central Electrochemical Research Institute, Karaikudi 623 006
Asis Datta, School of Life Sciences, Jawaharlal Nehru University, New Delhi 110 067
D. Balasubramanian, Centre for Cellular and Molecular Biology, Hyderabad 500 007
J. Chandrasekhar, Department of Organic Chemistry, Indian Institute of Science, Bangalore 560 012
R. Gadagkar, Centre for Ecological Sciences, Indian Institute of Science, Bangalore 560 012
K. N. Ganeshaiah, Department of Plant Genetics and Breeding, University of Agricultural Sciences, Bangalore 560 065
R. V. Hosur, Tata Institute of Fundamental Research, Bombay 400 005
N. V. Joshi, Centre for Ecological Sciences, Indian Institute of Science, Bangalore 560 012
Jyotirmoy Das, Indian Institute of Chemical Biology, Calcutta 700 032
C. C. Kartha, Division of Cellular and Molecular Cardiology, Sree Chitra Tirunal Institute for Medical Sciences and Technology, Thiruvananthapuram 695 011
S. S. Krishnamurthy, Inorganic and Physical Chemistry Department, Indian Institute of Science, Bangalore 560 012
S. C. Lakhotia, Department of Zoology, Banaras Hindu University, Varanasi 221 005
K. Muralidhar, Department of Zoology, Delhi University, Delhi 110 007
R. Narasimhan, CMC Limited, Bangalore 560 001
K. Neelakantan, ANURAG (Advanced Numerical Research and Analysis Group), Hyderabad 500 258
R. Nityananda, Raman Research Institute, Bangalore 560 080
T. Padmanabhan, Inter-University Centre for Astronomy and Astrophysics, Pune 411 007
B. R. Pai, Propulsion Division, National Aerospace Laboratories, Bangalore 560 017
T. J. Pandian, School of Biological Sciences, Madurai Kamaraj University, Madurai 625 021
Y. S. Rajan, Department of Science and Technology, New Delhi 110 016
M. Ramakrishnan, Geological Survey of India, Calcutta 700 016
A. V. Ramani, T. T. K. Pharma Ltd., Bangalore 560 025
S. Ramaswamy, Department of Physics, Indian Institute of Science, Bangalore 560 012
A. K. Raychaudhuri, Department of Physics, Indian Institute of Science, Bangalore 560 012
S. R. Shetye, Physical Oceanography Division, National Institute of Oceanography, Dona Paula, Goa 403 004
V. Siddhartha, 51 Bharati Nagar, New Delhi 110 003
V. Srinivas, School of Mathematics, Tata Institute of Fundamental Research, Bombay 400 005
G. Srinivasan, Raman Research Institute, Bangalore 560 080
R. Uma Shaanker, Department of Crop Physiology, University of Agricultural Sciences, Bangalore 560 065
K. S. Valdiya, Department of Geology, Kumaun University, Nainital 263 002
G. Venkataraman, Department of Physics, Sri Sathya Sai Institute of Higher Learning, Prasanthi Nilayam 515 134
M. Vidyasagar, Centre for Artificial Intelligence and Robotics, Bangalore 560 001
K. VijayRaghavan, TIFR Centre, Indian Institute of Science Campus, Bangalore 560 012

Current Science, founded in 1932, is published by the Current Science Association in collaboration with the Indian Academy of Sciences. The journal is also intended as a medium for communication and discussion of important issues that concern science and scientific activity. All articles published in *Current Science*, especially editorials, opinions and commentaries, letters and book reviews, are deemed to reflect the individual views of the authors and not the official points of view, either of the Current Science Association or of the Indian Academy of Sciences.

Working Committee, Current Science Association

P. Balaram, Indian Institute of Science, Bangalore
R. Chidambaram, Department of Atomic Energy, Bombay
V. K. Gaur, CSIR Centre for Mathematical Modelling and Computer Simulation, Bangalore
K. Kasturirangan, ISRO Satellite Centre, Bangalore
P. K. Kaw, Institute for Plasma Research, Gandhinagar
N. Kumar, Raman Research Institute, Bangalore (*Vice-President*)
N. V. Madhusudana, Raman Research Institute, Bangalore (*Secretary*)
A. Mani, Field Research Unit, Indian Institute of Tropical Meteorology, Bangalore (*Past-President*)
R. A. Mashelkar, National Chemical Laboratory, Pune
R. Narasimha, Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore
G. Padmanaban, Indian Institute of Science, Bangalore (*Vice-President*)
R. Ramachandran, The Institute of Mathematical Sciences, Madras
P. Rama Rao, Department of Science and Technology, New Delhi
S. Ramaseshan, Raman Research Institute, Bangalore (*Treasurer*)
M. S. Valiathan, Manipal Academy of Higher Education, Manipal (*President*)

Editorial Office

Editorial staff

Chandrika Ramesh, G. Madhavan, C. S. Ravi Kumar, M. S. Venugopal

Circulation and Accounts

Peter Jayaraj, Ranjini Mohan, B. Sethumani, Shanthi Bhaskar, B. K. Shivaramiah

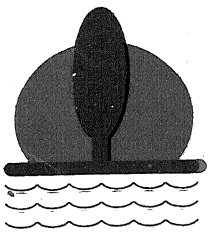
Annual subscription 1995

	Institutional	Personal
India	Rs 350	Rs 100
All other countries (AIR MAIL)	US $175	US $50

Single copies other than special issues: Rs 35/US $15

Editorial Office: CURRENT SCIENCE
C. V. Raman Avenue
P.B. 8001
Bangalore 560 080, India
Telephone: 91-80-3342310
Telex 845-2178 ACAD IN
Fax: 91-80-3346094

Copyright © 1995, Indian Academy of Sciences



IPS-11

July 28–Aug 02, 1996
Bangalore, India

Eleventh International Conference on Photochemical Conversion and Storage of Solar Energy (IPS-11)

Scientific Programme

The IPS conferences are held every two years with the aim of outlining the latest developments in the field of photochemical conversion and storage of solar energy. The eleventh conference of this kind will cover the following areas:

* Photoinduced electron transfer processes
* Photocatalysis in heterogeneous and homogeneous systems
* Photoelectrochemistry
* Photosynthesis in biological and biomimetic systems
* Quantum chemical studies of molecules on surfaces under the influence of light
* Thermodynamic aspects of quantum solar energy conversion
* Environmental applications of photocatalytic systems
* Materials for photoenergy conversion and photochemical engineering

These topics will be discussed in plenary and at invited lectures, workshops and poster sessions. Contributions from participants are encouraged. A book of abstracts will be distributed at the conference.

Location

The Eleventh International Conference on Photochemical Conversion and Storage of Solar Energy (IPS-11) will be held in the National Science Seminar Complex, Bangalore, India.

International Organizing Committee

E. Amouyal (France): C.A. Bignozzi (Italy): M. Calvin (USA); G. Calzaferri (Switzerland): Y. Cao (China); M. A. Fox (USA); M. Fujihira (Japan); D. Gust (USA); G. Hodes (Israel); V. Krishnan (India); J.-M. Lehn (France); N. S. Lewis (USA); S. E. Lindquist (Sweden); A. Nozik (USA); V. N. Parmon (Russia); E. Pelizzetti (Italy); Sir G. Porter (UK); N. Serpone (Canada); H. Tributsch (Germany); F. Wilkinson (UK); H. Yoneyama (Japan); F. Willig (Germany); K. I. Zamaraev (Russia)

Local Organizers

S. Umapathy and P. K. Das, Department of Inorganic and Physical Chemistry, Indian Institute of Science, Bangalore.

Second Circular

The second circular containing more details will be sent by October 1995. The Organizing Committee would appreciate your supplying this information to others who may be interested in attending the IPS-11.

Correspondence

The Secretariat IPS-11
Prof. V. Krishnan
Department of Inorganic and Physical Chemistry
Indian Institute of Science
Bangalore 560 012, India
Tel: 91-80-3092296
Fax: 91-80-3341683.

Founded: 1932

ISSN: 0011-3891

CURRENT SCIENCE

Volume 69 Number 5

10 September 1995

391 **In this issue**

NEWS

393 **Peer review under scrutiny**
C. C. Kartha

CORRESPONDENCE

395 **What NSU report leaves unanswered!**
Y. P. Joshi

396 **Proposal for a National Science School**
Oomen Abraham

SCIENTIFIC CORRESPONDENCE

397 **The ultimate constituents of matter**
K. Vasudeva Rao; *Reply by* K. V. L. Sarma

397 **Record of Upper Gondwana flora from Bairam–Belkher area of district Amravati, Maharashtra and district Betul, Madhya Pradesh**
A. K. Srivastava, P. D. Banubakode, V. M. Kale and G. V. Patil

400 **First report of the Late Permian sporangiate fructification *Nesowalesia* Pant from India**
Shaila Chandra and Kamal Jeet Singh

COMMENTARY

401 **Whither microbiology?**
Ramesh Maheshwari

406 **Finding *patterns* in nature's maze: An endless quest**
K. Chandrashekara

GENERAL ARTICLES

410 **INSAT Master Control Facility: An insight**
S. Rangarajan

415 **Vegetarianism: A nutritional appraisal**
K. T. Achaya

420 **Cursive English script recognition – A knowledge-based system**
P. V. S. Rao and T. M. Ajitha

REVIEW ARTICLE

426 **Wave-particle duality and Bohr's complementarity principle in quantum mechanics**
D. Sen, A. N. Basu and S. Sengupta

RESEARCH ARTICLE

434 **α-Turns in protein structures**
D. V. Nataraj, N. Srinivasan, R. Sowdhamini and C. Ramakrishnan

RESEARCH COMMUNICATIONS

448 **Task activation of the human brain studied in real time by functional magnetic resonance imaging**
N. R. Jagannathan and P. Raghunathan

452 **Radon precursory signals of Chamba earthquake**
H. S. Virk, Vivek Walia and Anand K. Sharma

454 **Palaeogeographic evolution of the Akkulam lake and its geotechnical implications in the urban development of Thiruvananthapuram**
K. Reghunadh, K. Sajan and S. Mohanakumar

458 **Characterizing intergeneric regenerants of protoplast fusion between *Hyoscyamus muticus* (Egyptian henbane) and *Atropa belladonna* (Indian sagangur)**
C. C. Giri and P. S. Ahuja

461 **Mechanism of wound healing induced by chitosan in streptozotocin diabetic rats**
A. Naseema, Pius S. Padayatti and C. S. Paulose

464 **Permeability of cuticle of *Helicoverpa armigera* (Hübner) larvae to deltamethrin**
A. P. Padma Kumari, A. Phokela and K. N. Mehrotra

BOOK REVIEWS

467 **Quark and Jaguar,** *reviewed by* A. N. Mitra

470 **Functional Dynamics of Phytophagous Insects,** *reviewed by* K. Chandrashekara and Raghavendra Gadagkar

471 **The State of Food and Agriculture 1993 and 1994,** *reviewed by* M. S. Swaminathan

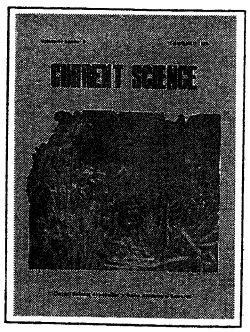
PERSONAL NEWS

473 **A visionary and courageous scientist,** *an obituary of Jonas Salk by* G. P. Talwar

474 **A doyen of Indian physicists,** *an obituary of S. D. Chatterjee by* D. N. Bose

HISTORICAL COMMENTARY AND NOTES

476 **A probe into the Calcutta earthquake of 1737**
Pitta Govinda Rao



COVER. Giant tubeworms, *Riftia pachyptila*, at hydrothermal vent (Courtesy of IFREMER, France). See page 401.

Indexed in CURRENT CONTENTS/GEOBASE/CHEMICAL ABSTRACTS

In this issue

Resurrecting microbiology

Sidney Brenner remarked some time ago that 1990 marked the demise of both biochemistry and communism. For good measure he categorized physiology as an extinct discipline (*Current Biology*, 1995, **5**, 332). Carl Woese recently suggested that 'if microbiology is not dead today, it definitely is moribund' (*Microbiol. Rev.*, 1994, **58**, 1). To the outsider, the present boom in biology would appear to contradict these remarkably harsh assessments on the state of many areas of the science. However, in the headlong rush to clone, sequence and sometimes express genes, most students of biology and bordering areas have gravitated to 'molecular biology'. Armed with a formidable array of cellular and molecular tools, identifying, cutting, pasting and modifying genes appears to be the most desirable activity for new students of molecular biology. Research in this area has advanced at a furious pace with an ever increasing number of companies providing ready-made kits to perform almost any experiment. In universities and institutions where researchers are trained, the unchecked growth of 'molecular biology' has been at the expense of several more traditional disciplines, resulting in a new generation of students unaware of the importance of microbiology, biochemistry and physiology. Are these murmurs of discontent only a case of sour grapes? After all, the molecular genetic approach has provided incisive solutions to many major biological issues and resolved many an impasse faced by the traditionalists. The successes of genome sequencing are already impressive, with the *Haemophilus influenzae* story being the most recent (*Science*, 1995, **269**, 496). Should scientific disciplines be given a graceful burial when their time is past? Or will life in many forgotten areas of biology, if nurtured, result in a new flowering in the future. Ramesh Maheshwari makes a strong and reasoned plea (**page 401**) for vigorously promoting microbiology as central to the study of all biology.

P.B.

Functional imaging

When Paul Lauterbur published the first NMR images of coaxial tubes (*Nature*, 1973, **242**, 190), few would have imagined that the new technique called 'zeugmatography' would herald the birth of magnetic resonance imaging (MRI). The prefix 'nuclear' was quietly dropped when the vast potential of the method in clinical medicine was realized. This undoubtedly enhanced the rate with which this amazing technique entered many areas of diagnostic medicine. MRI has progressed at an impressive pace helped by many fundamental advances in the practice of nuclear magnetic resonance; the introduction of high field superconducting magnets, the development of two-dimensional spectroscopy and the fusion of spectroscopy and imaging. Today MRI facilities exist at many centres in India, with the All India Institute of Medical Sciences (AIIMS) boasting of a Department of Nuclear Magnetic Resonance. The results of the first functional MRI experiments carried out on the 1.5 T whole body scanner at AIIMS are reported by Jagannathan and Raghunathan on **page 448**. The authors demonstrate the possibility of using conventional pulse sequences and equipment for functional imaging. Neuronal activity stimulated by specific physical tasks is shown to affect the magnetic resonance image. The possible applications in neurosciences are many. This will undoubtedly require a very close interaction between neuroscientists and spectroscopists.

P.B.

Patrons

Dr. S. K. Joshi
President, INSA

Dr. R. A. Mashelkar
Director General, CSIR

Dr. G. V. Satyabati
Director General, ICMR

Dr. C. R. Bhatia
Secretary, DBT

Dr. V. S. Ramamurthy
Secretary, DST

Advisory Committee

President
Prof. V. Ramalingaswami

Members
Prof. B. K. Bachhawat
Prof. L. S. N. Prasad
Prof. G. Padmanabhan
Prof. B. N. Dhawan
Prof. Indira Nath
Dr. Sandip K. Basu
Dr. Manju Sharma
Dr. Asish Mukherjee
Dr. V. P. Sharma
Prof. R. C. Mahajan
Prof. N. K. Ganguly
Dr. Lalit Kant
Dr. V. P. Kamboj
Prof. P. N. Tandon

Organising Committee

President
Prof. A. B. Chowdhury

Executive President
Prof. A. N. Bhaduri

Vice-Presidents
Prof. A. K. Hati
Prof. P. K. Ray
Dr. Santanu Kar
Dr. J. Das
Dr. D. Basu

Jt. Secretaries
Dr. M. K. Basu
Dr. H. K. Majumder

Treasurer
Dr. Dwijen Sarkar

Members
Prof. Asoke Ghosh
Prof. V. K. Vinayak
Prof. A. Nandy
Prof. A. Mukherjee
Prof. A. Bhattacharyya
Dr. Tuli Biswas
Dr. Mitali Chatterjee
Dr. Syamal Roy
Dr. Kawna Chatterjee
Shri Arun Banerjee
Smt. Debasree Mukherjee



SYMPOSIUM
on
U. N. BRAHMACHARI & PERSPECTIVE OF KALA-AZAR RESEARCH

Indian Institute of Chemical Biology
Calcutta 700 032

Date: 3–4 November 1995

TOPICS OF SCIENTIFIC SESSIONS

1. History of Kala-azar research and contributions of Professor Brahmachari
2. Clinical profile of the disease including PKDL
3. Epidemiology and transmission dynamics of the disease
4. Chemotherapeutic approach
5. Vector biology and control
6. Immunobiology of host–parasite interaction
7. Biochemistry and molecular biology of Leishmanial parasites
8. Pathophysiology of Leishmaniasis
9. Diagnosis
10. Control

Address for correspondence:

Dr. M. K. Basu/Dr. H. K. Majumder
Indian Institute of Chemical Biology
4, Raja S. C. Mullick Road, Calcutta 700 032
Phone: 473-6793/0492/5197, Fax: 473-5197/0284

CURRENT SCIENCE

Volume 69 Number 5 10 September 1995

NEWS

Peer review under scrutiny

The practice of seeking an expert's comments on submitted articles is as old as *Philosophical Transactions of the Royal Society* and is the central mechanism for assessment of the quality of manuscripts before publication in scientific journals. In recent years, a number of editorials and articles have addressed various issues related to the process of peer reviewing, reflecting the increasing interest on the subject[1-7]. Six years ago, thanks to the efforts of the editors of some of the leading medical journals, a conference on peer review was held in Chicago, Illinois[8]. The meeting focused attention on problems connected with manuscript assessment and helped to stimulate research into relevant aspects of peer review.

A second international congress on peer review in biomedical publications was held in Chicago, Illinois, in September 1993 and brought together editors and researchers from 20 different countries. *The Journal of the American Medical Association* (JAMA) has devoted one of its recent issues to publishing 26 articles selected from 110 papers presented during the conference[9].

The published papers discuss several interesting aspects such as effectiveness of peer review as a quality control mechanism, role of blinded reviews, author's views on editorial practices and cognitive aspects of manuscript assessment.

M. Abbey, M. D. Massey, S. Galandiuk and H. C. Polk (*The American Journal of Surgery*) determined what percentage of manuscripts rejected by their journal were subsequently published in other journals indexed by MEDLINE. They found that 62% of the manuscripts rejected by them remained unpublished in core medical journals. More than half of the articles, ultimately published, appeared in journals with a lower circulation rate. Their findings imply that review process effectively prevents publication of manuscripts of a lower quality.

The effect of peer review and editorial processes on the readability of articles has been measured by J. C. Roberts, R. H. Fletcher and S. W. Fletcher (*Annals of Internal Medicine*). Editorial practices were found to improve the readability of the articles as measured by indices such as the Gunning fog index and the Flesch reading ease score. Readers, mostly practising physicians, however did not agree with the peer assessment with respect to the importance of published articles. Articles which received high marks for importance were often judged as less relevant to clinical practice. This difference in opinion may be because the reviewers recommend articles containing new information and the clinicians are reluctant to accept unestablished findings.

Peer review is thought to be a good defence against scientific misconduct. Yet, numerous articles with fabricated or plagiarized or redundant (duplicate) data get published despite peer review. H. N. Nigg and G. Redulescu (*Bulletin of Environmental Contamination and Toxicology*) report four cases of misconduct encountered in their journal since 1990. In three cases, the authors belonged to institutions in India.

P. A. Rochar *et al.* of Mount Sinai Hospital confirm our suspicion that scientific articles published in journal supplements are generally of inferior quality when compared with those published in the parent journal. Their finding is based on assessment by blinded reviewers of the quality of articles on randomized control trials of drug therapies published in *American Journal of Cardiology*, *American Journal of Medicine* and *American Heart Journal.*

What is the impact of blinded peer review? D. N. Laband and M. J. Piette from the Department of Economics, Auburn University, drew a sample of 1051 articles in 28 economics journals published during 1984 and found that articles in journals using nonblinded peer review received lesser number of citations than those in journals using blinded peer review. Journals using nonblinded review appear to publish a large number of papers that should not have been published.

The basis for recommending blinded reviews is that many researchers have the apprehension that reviewers recommend and editors accept articles that originate from prestigious or better-known institutions and those submitted by authors who have more previous publications. J. M. Garfunkel and colleagues (*Journal of Paediatrics*) found in their analysis that there is no difference in the acceptance rate of major manuscripts from institutions of greater prestige and from other institutions. However, a significant correlation existed between institutional rank and recommendations or decisions to accept case reports and brief reports. M. Fisher, S. B. Friedman and B. Strauss (*Journal of Developmental and Behavioural Paediatrics*) also concluded from a randomized controlled trial that there was no significant difference in the scoring of manuscripts and recommendations when reviewers were either blinded or nonblinded to the previous publication record of the author.

Another form of bias suspected in handling of manuscripts is gender bias. J. R. Gilbert, E. S. Williams and D. G. Lundberg (JAMA) assessed whether the manuscripts received by the journal had different manuscript processing characteristics associated with the gender of the corresponding author, editor and reviewer. Participants in this retrospective cohort

study included 8 male and 5 female editors, 2452 male and 930 female reviewers and 1698 male and 462 female authors. Interestingly, the study revealed that gender differences exist in editor and reviewer characteristics. Female editors summarily rejected manuscripts at a higher rate than male editors ($p < 0.001$!). Male editors utilized male reviewers more often than female reviewers, while female editors used reviewers of both genders at equal rates. Female editors were assigned manuscripts with female corresponding authors at a higher rate. Male reviewers took significantly longer time to return the manuscripts than female reviewers. Gender of the reviewers and corresponding authors had no association with the recommendation made by the reviewers. There was no evidence of gender bias in the final acceptance rate of manuscripts.

M. Nylenna, P. Riis and Y. Karlsson (*Journal of Norwegian Medical Association*) studied the association between the characteristics of referees and their assessment of manuscripts, and also the influence of manuscript language. Experienced and young referees were found to provide better reviews than their older colleagues. In Scandinavian countries, an English version of a manuscript was more easily accepted than its national version, and the submission of an article in mother tongue appeared to be a handicap.

How do authors view editorial practices? The subject has been investigated by B. J. Sweitzer and D. J. Cullen (*Journal of Clinical Anaesthesia*). They ascertained from authors of unsolicited articles their views on the evaluation of their manuscripts. Opinion was sought on timeliness, quality, clarity, specificity and educational value of reviewer's comments as well as their satisfaction or dissatisfaction (because of delay, unclear comments, overlooking important data, tone of the review, lack of knowledge of the reviewer, etc.) with the reviews. Authors of accepted manuscripts gave a favourable rating for reviewers. Authors of articles in the category of 'accept with revision' were satisfied that the review process led to improvement in their manuscripts. Authors of articles in 'reject but resubmit' category believed that the review process improved subsequent preparation of manuscripts. Worst scores were provided by authors of 'reject outright' articles. Only 30% in this group responded and among them only 20% believed that they were helped by the review process.

Can the nature of manuscript assessment be improved to help editors in decision making? Editors often encounter situations when reviewers do not agree in their recommendations on a manuscript. On such occasions editors either use an arbitrary cut-off score to determine acceptance or use their discretion. Editors would be helped if the reviewers provide a structured assessment and there are objective criteria to measure the reliability and quality of reviews.

J. P. Kassirer and E. W. Campion (*New England Journal of Medicine*) in their paper discuss the cognitive aspects of manuscript assessment. They try to blunt the criticism that there are no rules governing peer review and that the process is arbitrary and subjective. They point out that the fundamental task of a manuscript reviewer is to 'detect and describe flaws' and that in the assessment of manuscripts there is a rejection threshold – a level at which the cumulative faults of a paper turn the scales towards rejection. Manuscript assessment also has a sensitivity and specificity. Wrong recommendations are unavoidable in the review process since even in the best of hands false positive and false negative results do occur. If this state has to change, the cognitive aspects of how a reviewer makes an assessment of manuscripts have to be better understood. J. P. Kassirer and E. W. Campion (*New England Journal of Medicine*) have categorized and listed the various aspects of a manuscript on which a reviewer is expected to comment. The categories include deficiencies of design, deficiencies in presentation, deficiencies in interpretation and assessment about the importance of research. Under each category flaws are listed. The list makes the reviewers' job in making objective assessment on articles easy.

M. K. Cho and L. A. Bero from the Institute for Health Policy Studies, California, have developed two instruments, one for measuring the methodologic quality and another to assess the clinical relevance of reports on clinical research. Both have been tested and found reliable, valid and applicable to a variety of research designs.

A grading instrument with a 14-point scoring system and a scoring guide to measure objectively the quality of reviews is provided by I. D. Feurer *et al.* (*Journal of Vascular Interventional Radiology*). The seven attributes in the system include timeliness, etiquette, completion of a grade sheet, and provision of (i) sectional reviews, (ii) supporting references, (iii) specific recommendations and (iv) new insights or perspectives. The authors have tested the scoring system and found it valuable to identify outstanding and weak reviewers.

Authors' misconduct receives considerable publicity while editorial misdeeds are unpublished. Dishonesty, favouritism, victimization and nondeclaration of conflict of interest are recognized as some of the unethical editorial practices. The true prevalence of editorial misdemeanour is unknown. D. G. Altman, I. Chalmers and A. Herxheimer from Oxford, England, draw attention to three cases of editorial misdeeds. One case is that the editor republished an article which appeared in another journal without consulting the author and then attacked it in an editorial, refusing a chance for the author to reply. Another case involved plagiarism of a commissioned review article. The third is an instance of not disclosing the editor's vested interest, which led to unfair assessment of a manuscript describing a randomized controlled drug trial. Altman and colleagues suggest that an International Medical Scientific Press Council be established to set a code of conduct for editors and to investigate when there is breach of prescribed conduct.

Journal publishers are concerned with the way in which authors select the journals for submission of their articles. Eric Frank from Emory University School of Medicine conducted a survey among the clinical and research faculty of Stanford University School of Medicine to evaluate authors' criteria. The most important factors for initial manuscript submission were journal prestige, the frequency with which a journal publishes related articles and the composition of the readership. Likelihood of acceptance, turn-around time, prior publication in a journal and acquaintance with the editor gained importance for subsequent submissions.

H. F. Judson of Stanford University draws attention to factors that threaten the peer review process. He writes: 'What was a high and interesting duty has become a wearisome chore'. Waning dedication and enthusiasm of the participants in the peer review process are likely to

lead to declining standards. He also cautions that peer review, which began as a measure to shield the autonomy of science from administrators and politicians, is threatened by internal politics such as rivalries between scientists and the likelihood of theft of ideas and insights by peer reviewers, who are often competitors of the authors of articles. Judson predicts that structural transformation in sciences would have its impact on peer review processes. Shortage of funds for research, intensification of competition, increasing pressure from funding agencies for research targeted at national needs, with emphasis on industrial and technological exploitation of research, are characteristics of the current transition. Profit motives determine the direction of research. Criteria for the evaluation of a research work are determined by non-scientists and the emphasis is on outcome and endpoints than on inputs. Yet another transformation is the development of electronic collaboration, which offers considerable options for interaction between editors, authors and readers. New procedures for manuscript assessment are expected to replace the current practices.

What did the participants feel after the congress? Attitude survey conducted after the conference revealed that a vast majority of respondents thought that peer review improves the quality of publication, that the journals should adopt uniform standards for peer review and that the editors should encourage research to establish baseline data on the prevalence of scientific fraud. Among the respondents, 95% said that conferences on editorial peer review should be held regularly in the future.

1. Bailar, J. C. and Patterson, K., *N. Engl. J. Med.*, 1985, **312**, 654–657.
2. Stossel, T. P., *N. Engl. J. Med.*, 1985, **312**, 658–659.
3. Smith, R., *Br. Med. J.*, 1988, **296**, 774–777.
4. Rennie, D. and Knoll, E., *Ann. Intern. Med.*, 1988, **109**, 181.
5. Relman, A. S., Angell, M., *N. Engl. J. Med.*, 1989, **321**, 827–829.
6. Mc Nutt, R. A., Evans, T., Fletcher, R. H., Fletcher, S. W., *J. Am. Med. Assoc.*, 1990, **263**, 1371–1376.
7. Fabiato, A., *Cardiovasc. Res.*, 1994, **28**, 1134–1139.
8. Guarding the guardians: Research on editorial peer review. Selected proceedings from the First International Congress of Peer Review in Biomedical Publications, 1990, **263**, 1317–1414.
9. *J. Am. Med. Assoc.*, 1994, **272**, 79–174.

C. C. Kartha, Sree Chitra Tirunal Institute for Medical Sciences and Technology, Thiruvananthapuram.

CORRESPONDENCE

What NSU report leaves unanswered!

This refers to the Special Section of *Current Science* (25 June) in which you have closed the debate on NSU. To me it appears that the debate has been ineffective and inconclusive. I was surprised and somewhat disappointed to note that you have not summarized in clear terms, with reasonable details, what your authors and readers have generally opined about the new university. In your 'In this issue' column, there appears a remark: 'there were few supporters of the NSU concept'. I just wonder if the term *few* is actually intended to mean *some* (equivalent to *a few*) or to imply *hardly any* (which it really does)! In the latter case, your authors' and readers' message is rather clear.

There are a few important questions that have been left unaddressed in the final report. Supporters of NSU and those actively engaged in its establishment should feel obliged to give serious thought to these. Some of them are:

1. NSU will be producing some 200 Ph D's every year. Where would they be expected to go? Would they accept to work in other Indian universities/colleges whose academic environment would hardly match that of the said university?

2. Will the phenomenon of brain-drain not apply to the trained scientists coming out of NSU? Will they be allowed to go abroad, since in a free and democratic country nobody can be stopped from searching for a better working place? Will we allow brain-drain, giving the argument that science is universal and that complete mobility must be granted to scientists so that they could work in a place of their liking? But then we have to remember that there are moments when we become specially conscious of our nationalism and start talking of things like Indian science.

3. The report states that NSU would also train foreign students. Why spend the available resources on foreign students? Why not train more and more Indian students? After all, the university would hardly be able to cater to the academic needs of our vast population. That means we would be obliged to concentrate our efforts fully on training as many Indians as possible.

I have a suggestion to make. All those who are involved in erecting NSU – a temple of academic excellence – should avoid accepting a faculty chair or any other position of importance on the completion of the task. In the end, they should hand over the institution to deserving people from within the country and from outside. This will reinforce the faith reposed on them by the general public and will also succeed in impressing upon the skeptics that NSU was not being created in their own interests but in the interest of the nation. Let us remember that great men, whose actions are motivated by a deep-rooted concern about the welfare of the society, gracefully leave a completed project to others to look after and move on to a newer venture.

Y. P. JOSHI

Department of Physics,
Banaras Hindu University,
Varanasi 221 005, India.

Proposal for a National Science School

Many responses to the proposal for a National Science University (NSU) and the report on 'University Education in Science' clearly show that this kind of a discussion was long overdue.

The objective of setting up the first three universities of the country – Bombay, Calcutta and Madras – in the middle of the nineteenth century was to train personnel to work for the administration, under British superiors. The Indian officers were expected only to obey orders. Because of this there was no emphasis on independent thinking or innovations in these universities. It was feared that these would lead to indiscipline. A university degree was (still is?) considered as a passport to a good job, to earn a decent living. In a job, one was expected only to respect the superiors, obey orders, perform duties and follow rules. At the time of independence there were nearly 20 universities, and now there are around 200. In spite of this phenomenal increase in the number, the university system still remains as a hangover of the colonial past. Only as recently as 1993–1994, introduction of 'vocational' components at the undergraduate level is realized as necessary. Science education in India suffers the most because of the lethargy to modify the system according to the changing times. There is an urgent need to reform science education radically at the school, college and university levels.

Increased political interferences in the functioning of the universities and the trend of appointing Vice-Chancellors on political considerations are the major reasons for their rapid deterioration. In initial appointments and in promotions, when 'ministerial approbations' can substitute for many years of hard work (in libraries and laboratories), how can meritorious and dedicated research workers enter our universities as faculty members and how can the existing faculty members sustain their interest in research. Something needs to be done (very urgently) to keep (politics and) politicians away from our universities.

The problems of research institutions are many. Most important among them are: (i) a career in science remaining the last resort for those who opted for the science stream and failed to make it to a professional course, (ii) confusion whether science is a 'profession' or a mere job and (iii) poor career prospects.

In all work environments there are three categories of people. The first category enjoys the job, gets involved in it completely and does not need any external motivation. These people are sincere and have a natural liking for their jobs. For them, job is like a favourite hobby and a mission. The second category considers the job only as a means to earn money; their involvement in the job is very little, and their enjoyment is elsewhere. They put in the least possible efforts necessary to continue in their jobs. The third category is interested in everything else except work. They are interested in money, promotions, authority, manipulating and controlling people, publicity, etc. These are the politicians of the workplace. They do not promote any cause other than their own career, for which they may be willing to do anything and everything. Our science establishments are also filled with these three categories. And the tragedy is that people belonging to the first category are very few.

To have more people belonging to the first category in our science establishments, we must start searching for students having a natural liking for science. This needs to be done at a very early stage. First, many regional science schools (RSS) need to be established (these schools can be similar to Sainik schools). This should be followed by a National Science School (NSS – similar to the National School of Sports). Admissions to RSS and NSS can be on the basis of a national talent search examination, at the end of lower primary school (LPS) and upper primary school (UPS) stages, respectively. The syllabus and teaching in the RSS and NSS must have greater emphasis on science. Once these students complete their studies in the NSS, they can be trained in a National Science Academy (NSA). The NSA can be similar to the National Defence Academy (NDA). After completing studies/training in the NSA, they can go to a National Science University (NSU). Starting an NSU without doing this ground work is very unlikely to bring about the desired changes.

Another ill plaguing our science establishments is the absence of any distinction between scientists and technicians. Research institutions are like any other government offices. Much in the same way as lower division clerks (LDCs) get promoted to Class I officers in the course of time, technicians also become scientists on the basis of seniority. This raises a basic question – Who is a scientist? Can anyone become a scientist? In defence services there are officers and other ranks. In medical profession there are doctors, nurses and pharmacists. In all these, the selection procedures are different, and the duties, responsibilities and pays are also different. The promotional avenues of other ranks are different from those of officers. Nurses and pharmacists do not become doctors on the basis of seniority. Why is there no such distinction in science? Excepting ICAR, where there is an ASRB (Agricultural Scientists Recruitment Board), why is there no Scientists' Recruitment Board?

Lastly, will this NSU become a 'super IIT', from where scores of young Ph D's and faculty members would migrate to greener pastures? In that case, what purpose will it serve the country? We must stop producing professionally qualified manpower for other countries. There has to be at least a rough estimate on our need for trained/skilled manpower in different areas, and only somewhere near that many people should be trained. The service conditions of scientists need to be made attractive so that best students opt for science education and a career in science.

OOMEN ABRAHAM

341 Rasta Peth,
Pune 411 011, India.

The ultimate constituents of matter

For over 25 centuries, human curiosity about the ultimate constituents of matter in the universe has remained without a satisfactory solution. Ancient scientists recognized five fundamental elements, viz. earth, water, fire, air and ether. Yet, most philosophers believed that there must be a single kind of homogeneous material that constitutes the universe. Much of the philosophical thinking is guided by this axiom. On this point the schism of intellectuals into the 'two cultures' is not of recent origin. With the progress of science, the count of chemical elements grew from five to over a hundred. Subatomic particles too are not of one kind. However, philosophers as well as some scientists expect to find homogeneity and uniformity in matter in its minutest form. This belief seems to be the main prop for the 'modern scientific' interpretation of the philosophy of the *Upanishads*. As K. V. L. Sarma has reviewed admirably (*Curr. Sci.*, 1995, **68**, 605) the current state of knowledge, we now have a standard model (SM) of six quarks and six leptons, besides the gauge bosons, the transmitters of force.

Although Sarma does not discuss any 'philosophical' implications of the SM, any reader of the review will wonder if in ultimate analysis we might reach a paradigm of uniformity and homogeneity. Shall we conclude that philosophers' paradigm is only an idol of the tribe (*idola tribus*), to use Francis Bacon's terminology? In other words, can we conclude that uniformity of matter in its minutest form or in its primeval state at the Big Bang is a theoretical impossibility? If so, what could probably have guided ancient philosophers to their belief on the nature of matter?

K. VASUDEVA RAO

44 Anandavan,
A-6, Paschim Vihar,
New Delhi 110 063, India.

K. V. L. Sarma replies:

With regard to the comment of K. Vasudeva Rao, the speculation that there is some homogeneous stuff underlying the physical universe is perhaps quite natural and is indeed at the heart of the reductionist approach. History has repeatedly shown that things become simple when we go to the next substratum – molecules, atoms, nuclei, nucleons,... Maybe this is the only monotonous way Nature works, as we go on unravelling the successive onion layers.

There have been many attempts to construct composite models viewing the quarks and leptons, as well as the gauge bosons, as composites of the same fundamental objects, variously called as 'preons', 'rishons', 'subquarks'.... These models are unsatisfactory because proliferation is still present in them although at a different level (postulating many kinds of preons and hyper colour forces to bind them), and they do not lend themselves to any meaningful experimental tests. However, as of now, there is no experimental evidence to suggest that quarks or leptons have any structure, leave alone a common one.

K. V. L. SARMA

Tata Institute of Fundamental Research,
Homi Bhabha Road,
Bombay 400 005, India.

Record of Upper Gondwana flora from Bairam–Belkher area of district Amravati, Maharashtra and district Betul, Madhya Pradesh

This note records the plant fossils from Gondwana succession of Bairam (21°22′–77°37′)–Belkher (21°22′–77°31′) area, falling partly in Amravati district of Maharashtra and partly in Betul district of Madhya Pradesh. The fossils are represented by *Ptilophyllum*, *Sphenopteris*, *Cladophlebis*, *Taeniopteris*, *Matonidium*, *Pagiophyllum*, *Elatocladus* and *Araucarites*.

In regional set-up the area is represented by the Archaean Gneisses, Gondwanas, Lametas and Deccan Trap[1,2] (Figure 1, Table 1); however, the Archaean Gneisses is not exposed in the area. The Gondwana rocks are exposed in the form of isolated patches in the widespread Deccan Trap. It is about 100 m sequence represented by clastic sediments of arenaceous and argillaceous nature, having variable thicknesses (Figure 2). Pinching of the strata is frequently noticed in the beds of sandstones and clays in the upper part of the sequence. Clay is noticed to be of variable colours, viz. grey, dark grey, red, violet and brown to yellowish brown. Incidentally, grey to dark grey clays have been found productive for megafossil remains.

The age of the sequence is still a matter of debate. On the basis of lithological similarities, stratigraphic position, absence of fossils, it was earlier considered equivalent to Kamthi and (?) Mahadeva[1,3]. In the same way, Pachmari and Malari of Upper Triassic have also been correlated[4]. However, the age and correlation proposed, so far, were very casual and no macropalaeontological criteria have been taken into consideration. In recent years, the sequence has been ranked Neocomian–Aptian[5] on the basis of palynology.

The present assemblage is dominated by species of *Ptilophyllum* though *Taeniopteris* and *Matonidium* are very common, followed by *Sphenopteris* and *Cladophlebis* (Figure 3). The rest are poorly recorded. The recorded assemblage is comparable with Early Cretaceous assemblage of the Upper Gondwana

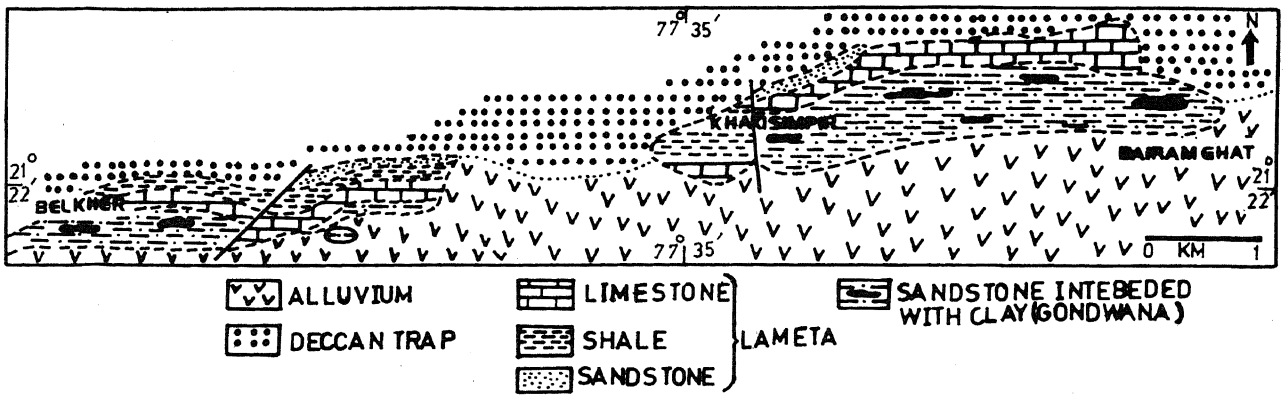


Figure 1. Geological map of the area (after ref. 2).

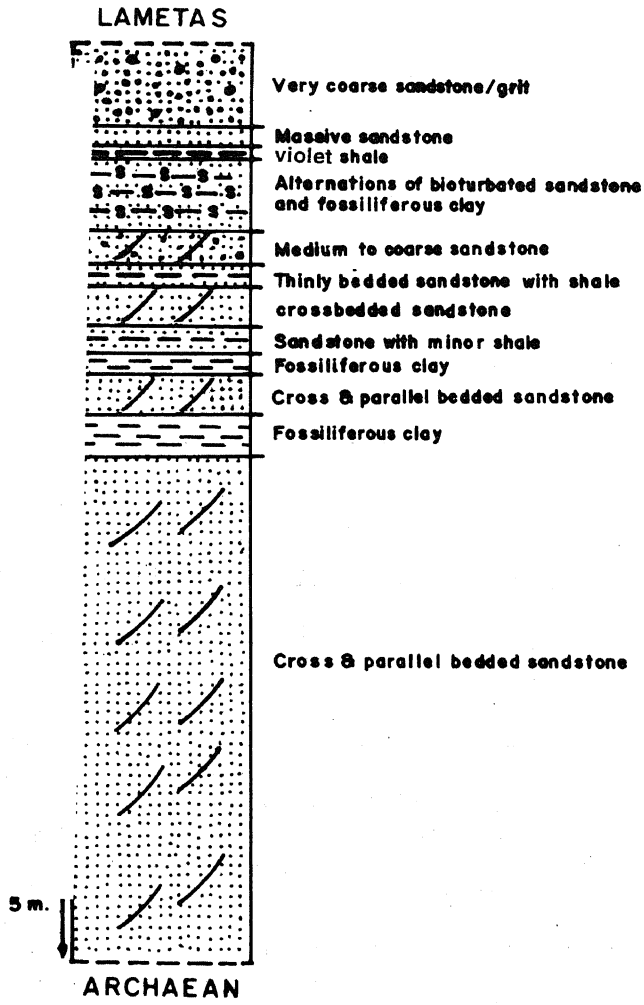


Figure 2. Lithological succession of the Gondwana sediments.

Formations of peninsular India (Table 2). However, the dominance of *Ptilophyllum* and *Taeniopteris* makes it more closer to Dubrajpur, Gollapalli and Jabalpur flora, which are Late Jurassic to Early Cretaceous in age[1,8–12]. The frequent occurrence of *Matonidium* is also comparable with Early Cretaceous of Himmatnagar[11,13]. The palynological evidences are also suppor-

Table 1. Generalized stratigraphy of the area

Age	Formation/Group	Lithology
Quaternary		Soil and alluvium
	Unconformity	
Miocene to Late Cretaceous	Deccan Trap	Nonporphyritic to porphyritic basalt
	Unconformity	
Late Cretaceous	Lametas	Shale, limestone and sandstone
	Disconformity	
Early Cretaceous	Upper Gondwana	Clay, shale, sandstone and conglomerate
	Unconformity	
Archaean		Quartzofeldspathic gneiss

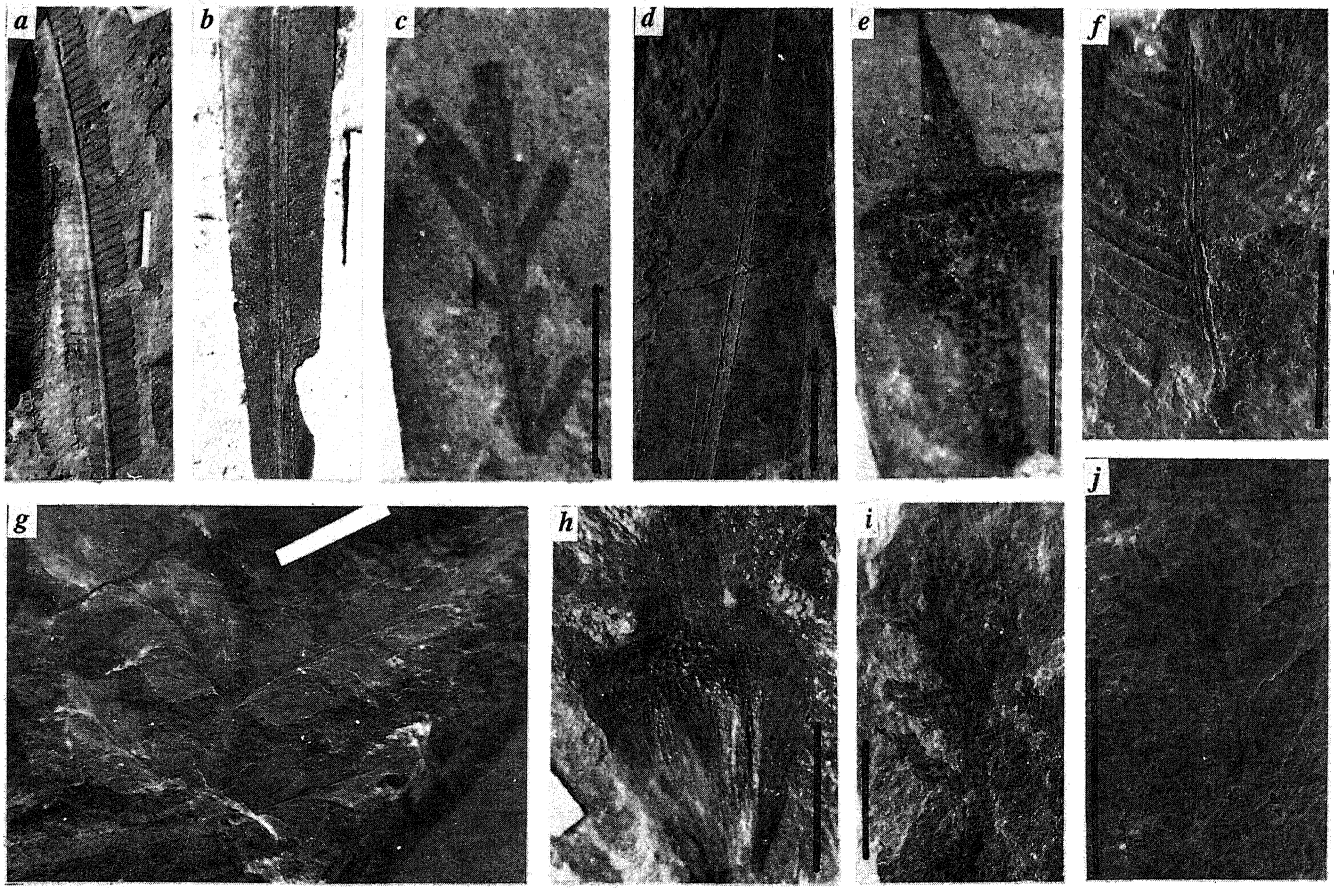


Figure 3. Photographs showing: *a*, *Ptilophyllum*; *b*, *Taeniopteris*; *c*, *Elatocladus*; *d*, *Taeniopteris*; *e*, *Araucarites scale*; *f*, *Matonidium*; *g*, *Sphenopteris*; *h*, *Araucarites*; *i*, *Pagiophyllum*; *j*, *Cladophlebis* (scale: 1 cm).

tive of Early Cretaceous age of the sequence[5]. Hence, the age and correlation proposed so far by the earlier workers is rejected and Early Cretaceous age is assigned.

The climate was warm and humid, as represented by luxuriant growth of Bennettitales and Cycadales[14]. This is also supported by pollen and spores reported from the area[5]. The dominance of cycadophytes over the conifers is also suggestive of tropical to subtropical climate as proposed for Early Cretaceous megafossil assemblage of Dubrajpur[8].

The samples are preserved in the P.G.

Table 2. Comparison of the Amravati plant fossil assemblage with that of other localities known from peninsular India

Genus/Locality	Present locality	Gardeshwar[6] (Early Cret.)	Gangapur[7] (Early Cret.)	Dubrajpur[8] (Late Jura. to Early Cret.)	Gollapalli[9] (Early Cret.)	Jabalpur[1,10,12] (Late Jura. to Early Cret.)	Umia[1,10,12] (Early Cret.)
Ptilophyllum	Abundant	—	+	Abundant	+	Very Common	+
Taeniopteris	Common	—	+	+	Abundant	Abundant	+
Cladophlebis	+	+	+	+	+	Common	+
Sphenopteris	+	+	—	—	+	+	—
Pagiophyllum	+	Abundant	+	+	+	+	+
Elatocladus	+	—	Abundant	—	+	+	+
Matonidium	Common	—	—	—	—	—	—
Araucarites	+	+	—	—	+	Common	+

+, Present
—, Absent.

Department of Geology, Amravati University, Amravati, Maharashtra, bearing numbers AU/G/R-1 to AU/G/R-10.

1. Pascoe, E. H., *A Manual of Geology of India and Burma*, Govt of India Publ. II, 1959, 3rd edn, pp. 1056–1057.
2. Gadewar, A. J. and Sukhtankar, R. K., *Gond. Geol. Mag.* 1990, **3**, 25–31.
3. Adyalkar, P. G., *Palaeogeography, Frame Work of Sedimentation and Groundwater Potentiality of the Upland and Alluvial Valley of India*, Govt of Maharashtra Publ., 1975, pp. 48–49.
4. Bhusari, A. G., *Min. Res.*, 1979, **IX**, 8–10.
5. Kumar, P., *Palaeobotany*, 1990, **39**, 381–386.
6. Bose, M. N., Banerjee, J. and Jana, B. N., *Curr. Trends Life Sci.*, 1979, **X**, 483–498.
7. Bose, M. N., Kutty, T. S. and Meshwari, H. K., *Palaeobotany*, 1982, **30**, 121–142.
8. Banerjee, J., *Palaeobotany*, 1989, **38**, 122–130.
9. Pandya, N. and Sukh-Dev, *Palaeobotany*, 1989, **38**, 147–154.
10. Pascoe, E. H., *J. Geol. Soc. India Mem.*, 1991, **21**, 50–67.
11. Shah, S. C., Singh, G. and Shastry, M. V. A., *J. Geol. Soc. India Mem.*, 1991, **21**, 97–112.
12. Bose, M. N., Proceedings of Symposium on Floristic Stratigraphy of Gondwanaland, Special Session, BSIP, Lucknow, 1966, pp. 143–154.
13. Banerjee, J., Jana, B. and Bose, M. N., *Curr. Trends Life Sci.*, 1979, **X**, 463–473.
14. Ramanujam, C. G. K., *Indian Gymnosperm in Time and Space*, Today & Tomorrow's Publ., New Delhi, 1976, p. 15.

ACKNOWLEDGEMENTS. We thank Shri P. N. Jagtap, Retd. Director, G.S.D.A., Amravati, for useful discussions. Dr Shaila Chandra, Dr J. Banerjee, Dr Ashwini K. Srivastava and Dr S. C. Srivastava from B.S.I.P., Lucknow, are thankfully acknowledged for identification of samples. Financial support to one of us (AKS) in the form of Minor Research Project (unassigned grant, UGC), Amravati University, is thankfully acknowledged.

A. K. SRIVASTAVA
P. D. BANUBAKODE
V. M. KALE
G. V. PATIL*

P. G. Department of Geology,
**Vice Chancellor,*
Amravati University,
Amravati 444 602, India.

First report of the Late Permian sporangiate fructification *Nesowalesia* Pant from India

Nesowalesia edwardsii, a Late Permian, sporangiate, monotypic fructification, is instituted by Pant[1] for concave disc-like organs from Newcastle, New South Wales, Australia. *Arberiella* type of sporangia are found attached on the concave side of the disc, while the convex side is smooth. Later on Rigby and Chandra[2] presented additional information concerning *Nesowalesia edwardsii* Pant on the basis of carbonized specimens found from the Late Permian Illawara Coal Measures of the Western Sydney Basin, New South Wales. They also found well-preserved taeniate bisaccate pollen *Striatopodocarpites solitus* (Bharadwaj and Saluja) Foster[3] from the *Arberiella* type of sporangia obtained from a thick crust of carbonaceous concave discs of *Nesowalesia edwardsii*. Rigby and Chandra[2] also proposed a peltate disc type of restoration for this Australian male fructification.

Two such discs having *Arberiella* type of sporangia are recorded for the first time from Late Permian bed exposed in Hinjrida Ghati near Handapa village, district Angul, Orissa. The concave discs are 1.5 cm in diameter and bear a number of *Arberiella* type of sporangia. Since the specimens are in the form of impressions, further studies are not possible. A close comparison between Indian and Australian specimens is also not possible as the

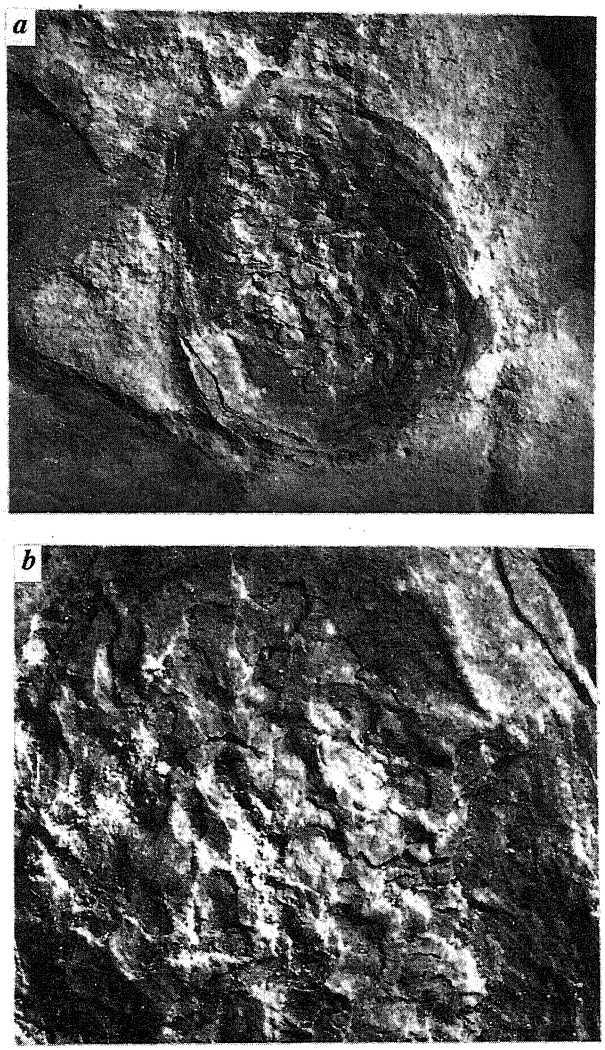


Figure 1. *a*, *Nesowalesia pantii* sp. nov. Specimen showing a concave disc with sporangia. Holotype specimen no. 36869 (×3); *b*, Enlargement of a portion of specimen in *a* showing numerous sporangia (×8).

mode of preservation in the two is different, the Indian ones being impressions while the Australian ones are carbonized. Indian specimens from Handapa are designated as a new species *Nesowalesia pantii* (holotype specimen number 36869, Birbal Sahni Institute of Palaeobotany, Lucknow) with a diagnosis 'concave disc-shaped organs with *Arberiella* type of sporangia' (Figure 1). The Indian specimens closely resemble the specimen described by Pant[4], photographed as Figure 7 on Plate 21.

Report of *Nesowalesia pantii* from Handapa bed is very significant as it is the first report of this sporangiate fructification from India.

1. Pant, D. D., *J. Indian Bot. Soc.*, 1977, **56**, 1–23.
2. Rigby, J. F. and Chandra, S., *Rev. Palaeobot. Palynol.*, 1990, **65**, 175–178.
3. Foster, C. B., *Publ. Geol. Surv. Queensl.* 372, *Palaeontol. Pap.*, 1979, **45**, 1–237.
4. Pant, D. D., *Bull. Br. Mus. Nat. His. Geol.*, 1958, **3**, 127–175.

SHAILA CHANDRA
KAMAL JEET SINGH

Birbal Sahni Institute of Palaeobotany,
53, University Road,
Lucknow 226 007, India.

COMMENTARY

Whither microbiology?

Ramesh Maheshwari

A distinguished American microbiologist, C. R. Woese, has expressed deep concern at the present state of microbiology. He is of the opinion that all biology should be reorganized around microbiology. The merit of this view is examined. Some examples are given to illustrate the contributions of microorganisms to the understanding of biological principles and formulation of concepts. A possible area of microbiological research is identified which could become the focus of activity involving many young scientists in our country and contribute to the vitalization of microbiology.

In a recent commentary on the state of microbiology, Carl Woese expressed concern thus[1]: 'University microbiology departments are withering. We are eliminating them as entities in their own right, being demoted to ineffectual

"programs" within departments whose names typically begin with "molecular" or "cellular".... In those that remain as microbiology departments in name, the emphasis in both teaching and research is shifting away from microbiology *per se*....' Increasingly, microbiology in these departments is becoming confined to that of growing of microorganisms for 'cloning, sequencing and expressing genes'. Elsewhere[2], Woese says: 'Microbiology is inadequately represented in the funding agencies. Students (and professors) genuinely interested in microbiology are increasingly hard to find. Microbiologists are afraid to teach real microbiology any more.... If microbiology isn't dead today, it definitely is moribund.' One may dismiss these criticisms as prompted from some occupational frustrations, such as the lack of sufficient funds for work, or lack of success in scientific research or the desire to attract attention. I am certain that none of these apply in the case of Woese.

Many biologists would recall that in the last decade somebody had pulled off a coup in biology. We had learned that all living organisms can be classified into two superkingdoms based on fundamental differences in cellular features: the prokaryotes (cells without a nucleus) and the eukaryotes (cells with a true nucleus). The tree of life had two main stems – one stem prokaryotic comprising the bacteria, and the other eukaryotic comprising plants and animals, the latter derived from the former. Woese's studies brought a fundamental revision of this picture. Examining the relationship between prokaryotic microorganisms by comparing the sequences of ribosomal RNA, Woese made the surprising discovery that the world of prokaryotes comprises unusual microorganisms, very different from the typical bacteria found in soil and water. (Imagine Woese's disbelief when, for example, a methane-producing microorganism which had all along been considered a bacterium turned out to be not a bacterium at all!) These microorganisms, which Woese named the archaebacteria (now called the Archaea), are found in unusual places such as hot springs, rumen of cattle, salt pans and brines. They not only differ from most bacteria in sequence of ribosomal RNA but also in biochemical features, such as the absence of muramic acid in cell wall and the presence of ether-linked membrane lipids. The archaebacteria are closer to eukaryotes than to typical bacteria. Their discovery showed that life on earth consists of not two but three primary domains: the Bacteria, the Archaea and the Eucarya. The new tree of life (Figure 1) is a framework for analysing the evolution of all life on earth. Woese placed the study of microbes in an ecological and evolutionary context. In recognition of his contributions, a hyperthermophilic archaeon, *Pyrococcus woesei*, which can grow at near 100°C, was named after him.

The concern expressed by this dedicated and distinguished microbiologist merits serious consideration. Woese contends that all of biology should be reorganized around microbiology. Surely, nobody will deny that microorganisms are important. Microbes cause diseases, make soil fertile, they allow cattle to digest grass, they clean up the environment, and they are used in the manufacture of antibiotics and alcohol. Beer and wine making, which involves alcohol fermentation by the yeasts, are very profitable industries! But, is the general, medical or industrial importance of microorganisms the compelling reason for biology to be reorganized around microbiology? Is the study of microorganisms crucial for learning and formulating biological mechanisms? Although I have no pretensions of being a real microbiologist, I believe that this narrative would demonstrate, at least approximately, the importance of microbiology in *all* biology. The report identifies a possible area of activity which could involve many microbiologists in our country and contribute to the vitalization of microbiology.

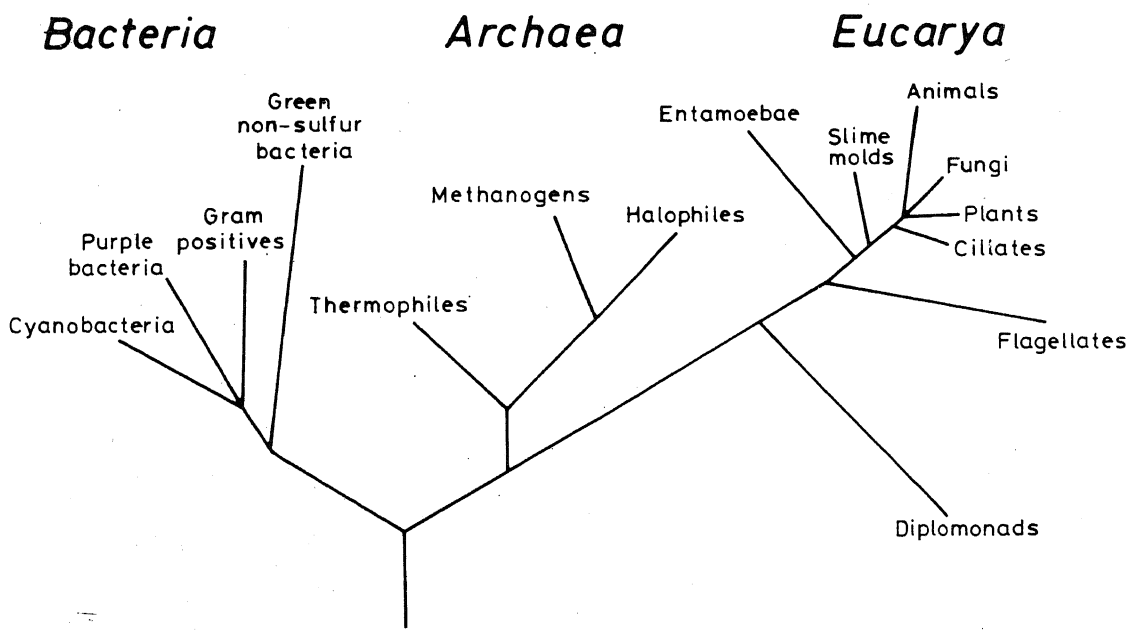


Figure 1. Universal phylogenetic tree determined from ribosomal RNA sequence comparisons showing the three domains, Archaea, Bacteria and Eucarya (based on ref. 2).

Teaching of microbiology

Microbial growth. We begin by the gesture of putting a few cells of a bacterium or yeast in a nutrient medium and measure cell numbers at regular intervals. Let us plot the logarithms of the numbers of cells as ordinate against time. The population displays a characteristic pattern of increase, the S-shaped sigmoid curve. We observe that after a short lag, the amount of living matter increases not in direct proportion to time, but according to geometric progression with time (exponential growth). In other words, the living matter multiplies itself by a constant factor in each successive unit of time. We learn that life is an autocatalytic process. However, the growth which started unrestrictedly stops abruptly when the population runs out of resources[3]. In this planet with dwindling supply of resources, the microbial growth curve has an obvious message for the humans.

Enrichment culture. Next, we prepare a synthetic medium containing a carbon source and mineral salts but lacking in nitrogen. We inoculate this medium with a pinch of garden soil and incubate it at room temperature. We note that microbes with specific metabolic properties develop in this medium. In this particular case, the medium selectively allows growth (enrichment culture) of microorganisms that can fix atmospheric nitrogen to the exclusion of numerous other microbial types present in the soil[3]. This simple exercise strikingly demonstrates the principle of natural selection in just a matter of few hours!

Microbial metabolism. We turn our attention to microbial metabolism and get acquainted with the sequence of reaction steps beginning from glucose and leading to the formation of ethyl alcohol, lactic acid, acetic acid or glycerol, depending upon the microbial species[3]. It suffices to note here that through a comparative, penetrating study of the dissimilatory pathways of microbial metabolism distributed among numerous types of bacteria, the great Dutch microbiologist A. J. Kluyver discerned that the recurring patterns in the variety of metabolic pathways are but few elementary reactions such as transhydrogenation, transphosphorylation, decarboxylation, etc[4]. From this was developed the concept of unity of biochemistry. A certain chemical conversion occupies a predominant position in a particular microorganism. For example, the conversion of pyruvate into acetaldehyde by pyruvate decarboxylase is predominant in yeasts which carry out alcoholic fermentation but is absent in lactic acid bacteria which carry out lactic acid fermentation.

Regulation. The existence of regulatory mechanism can be traced to the discovery in 1861 by Louis Pasteur, who observed that despite the lower consumption of glucose, more yeast is formed in the presence of air than in its absence (Pasteur effect)[5]. It stimulated intensive research on identification of the control sites in catabolism of glucose. Later, the observations that addition of certain amino acids to bacterial cultures shuts off the entire sequence of reactions, leading to the formation of the amino acid, led to the formalization of the concept of feedback inhibition[3]. The phenomena of induced synthesis of enzymes and the existence of specific membrane transport protein (permease) were discovered from studies dealing with growth and nutrition of bacteria[3].

Photosynthesis. The green cover on the earth has inspired many authors to comment that it is because of the remarkable process of photosynthesis invented by the plant chloroplast that man obtains food for himself and his cattle, fibre for his clothing, wood for fuel, his boat and his flute, and the paper for writing his poetry or scribbling his philosophy. A cover-up here? 'Chloroplasts did not evolve as a part of the plant cell; they were acquired by the (evolving) plant cell from the bacterial world through endosymbiosis', says Woese. 'In other words', he says, 'it is the bacteria that have invented photosynthesis, and plant chloroplasts are merely domesticated cyanobacteria'. The cyanobacteria are the most self-sufficient organisms in the biosphere. They not only carry out photosynthesis but also fix gaseous nitrogen from the atmosphere into biologically useful nitrogen compounds – a biochemistry which a vast majority of organisms are incapable of.

'We owe a debt of gratitude to the microorganisms whose peculiarities have provided a general insight into the understanding of photosynthesis', said van Niel[4]. What an ingenious use Engelmann (1880) made of certain motile bacteria as sensitive indicators of oxygen to demonstrate the direct connection between oxygen evolution and light absorption by chloroplasts in *Spirogyra* filaments! By illuminating the alga with light from a prism, he was able to determine the action spectrum of photosynthesis[4].

It was on the basis of the knowledge of a variety of microorganisms and the application of the concept of the unity of biochemistry developed by his master, A. J. Kluyver, and most importantly on the basis of the observations that (a) several types of bacteria can synthesize their cell material from CO_2 without the benefit of radiant energy, (b) the purple sulphur bacteria can grow anaerobically in a minimal medium supplemented with sulphide and bicarbonate, provided they are exposed to light, and (c) photosynthetic activity of the purple bacteria is not accompanied by evolution of O_2, that van Niel was led to conceptualize photosynthesis as a photochemical oxidation–reduction process and formulate his famous equation[4]:

$$\underset{\text{H donor}}{2H_2A} + \underset{\text{H acceptor}}{CO_2} \rightarrow \underset{\text{Oxidized H donor}}{2A} + \underset{\text{Reduced acceptor}}{(CH_2O)} + H_2O.$$

What this equation implied was that plant photosynthesis and bacterial photosynthesis are fundamentally similar light-dependent reactions. The difference is in the hydrogen donors they employ – being H_2S in the case of sulphur bacteria and H_2O in the case of the green plants. A direct consequence of the above-mentioned general formulation was the prediction that O_2 produced during photosynthesis by the green plants must come directly from water and not from CO_2.

Chemosynthesis. It was through the study of the life-style of a number of bacteria (nitrifying bacteria, sulphur bacteria, iron bacteria, etc.) found in nature under aerobic and anaerobic conditions that it was discerned that in addition to photosynthesis there is another process, chemosynthesis. In chemosynthesis the organic matter of the organism is built from CO_2 as the sole source of carbon without the supply of the radiant energy. As was first shown by the Russian microbiologist Winogradsky, a variety of bacteria can derive energy for growth from the oxidation of inorganic compounds[3,6]. That chemical energy can be derived from the biological oxidation of hydrogen, sulphide, sulphur, ammonia, carbon monoxide, ferrous, etc., was a revolutionary idea. It showed the apparently diverse ways in which microorganisms have solved the problem of

providing for their indispensable energy requirements.

Chemosynthesis had been thought to be a rare phenomenon until 1977, when scientists took a submersible at a depth of some 3000 m in the Pacific ocean. They discovered jets of superheated black water ejecting from the ocean floor and mixing with the cold sea water[7,8]. They were astonished to see a luxuriant fauna thriving in the warm water, in the vicinity of the hydrothermal vents (Figure 2). What is the food source for these new forms of animal life in the permanently dark environment? Pioneering work by Holger Jannasch showed that the ecosystems of animal populations are supported by microbial chemosynthesis rather than photosynthesis[7]. A variety of chemoautotrophic bacteria live in dense populations in the animals. The bacteria obtain energy by oxidizing geothermally provided reduced sulphur compounds and fix inorganic carbon through the Calvin–Benson pathway (Figure 3). The organic carbon is taken up by the animal hosts. The hosts are among the few animals on earth that are not fed directly by photosynthetically fixed carbon.

Thermophilic microorganisms. In the last decade another revelation occurred in microbiology. Karl Stetter isolated anaerobic archaea from sulphur-rich, marine volcanic environments that are able to grow at extremes of temperatures[9]. The upper temperature limit of life was pushed up to 110°C. Microbiologists are competing to see who can find microorganisms that beat this record!

The diversity of hyperthermophilic archaea has, in turn, started the race to discover the enzyme which has the highest degree of thermostability. Incredible it may seem, but an α-amylase from *Pyrococcus furiosus* retains enzyme activity even after autoclaving at 120°C for 5 h (ref. 10)! These microorganisms are the focus of the emerging area of biocatalysis at extreme temperatures and for biotechnological opportunities.

Early life. And yet another revelation. Thomas Gold predicted that microbial life is widespread at depth in the crust of the Earth[11]. He believes that in mass and volume the subcrustal life may be comparable with all surface life. Some microbiologists have sampled fluids from borewells drilled some 4000 m into the granitic rock or into the sea bed and have reported finding thermophilic anaerobic bacteria[12,13]! These surprising findings have fundamentally altered the views on the physical setting under which life arose and of the nature of early life on earth. Life probably arose in a hot environment. The ancestor may have been a chemosynthetic thermophile[14].



Figure 2. Giant tubeworms, *Riftia pachyptila*, at hydrothermal vent (courtesy of IFREMER, France).

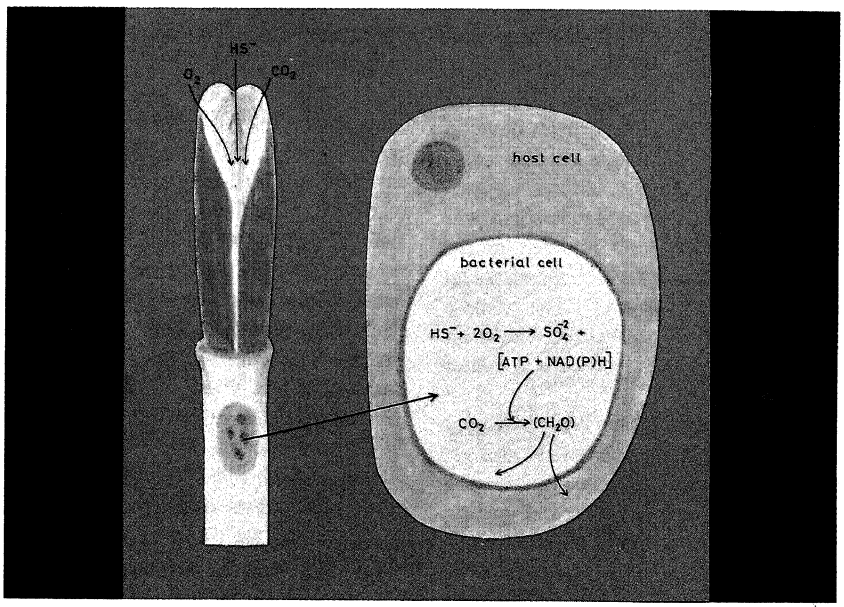


Figure 3. Schematic representation of symbiotic relationship between *Riftia* and chemosynthetic bacteria. Hydrogen sulphide serves as the electron donor, or energy source. The energy is produced when hydrogen sulphide and oxygen are combined and it is used for fixation of CO_2 (based on refs 7, 8).

Microbial diversity. 'The essence of the microbial world lies in its diversity', says Woese[2]. An innovative application of the techniques of molecular biology has been made to estimate how much of microbial diversity had escaped detection by conventional pure-culture technique in an

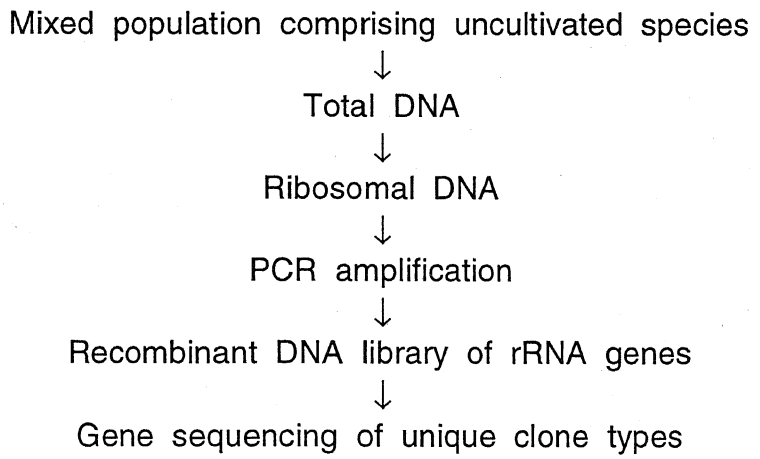


Figure 4. Sampling of microbial populations through isolation and sequencing of ribosomal RNA genes directly from the environment[15].

already well-studied place. Norman Pace took a census based on unique ribosomal RNA genes extracted and cloned directly from a niche (Figure 4)[15]. He demonstrated that more species of archaea were present than suspected.

It has been estimated that of the total number of species of bacteria (40,000) and the fungi (150,000), only about one-tenth have been discovered[16]. The fascination for microbial diversity comes from their strangeness and their exotic habitats. J. R. Postgate[17] gives a fascinating account of the specialized microbes which live at the bottom of deep oceans, in the mud in the estuaries of the rivers, in sewage treatment plants, in the interior of compost heaps, in the rumen of cattle, in near-boiling sulphuric acid conditions, in alkaline environments equivalent to a solution of caustic soda, in strongly salty places, in root nodules of leguminous plants, etc. A generalization that can be made from the results of microbiological sampling of diverse habitats is that no matter how harsh the conditions are in nature, so long as liquid water is available, microbial life will be found.

The importance of studying diversity has been beautifully stated by David Perkins[18]:

'Knowledge of a flute or a kettledrum is not sufficient to understand all the other instruments in a symphony orchestra or to predict their characteristics. Nor is a knowledge of a single species, however complete, adequate for understanding diverse species. Diversity of research organisms in the laboratory must at least dimly reflect the diversity of species in nature if the scope and beauty of evolutionary improvizations are to be appreciated and the genetic manipulation that brought them about are to be understood.'

Some other contributions of microbes. Before this 'refresher' course is closed, let us cast a final breezy look. Other examples where microorganisms contributed to new discoveries are: gene regulatory proteins[19], simplest light-driven proton pumps in halophilic bacteria[19], restriction modification system[19], random and directed mutations[20], checkpoints in cell cycle[21]. Lest we forget, the success in obtaining nutritional mutants in the fungus *Neurospora crassa* had led to the discovery of the relationship between genes and enzymes. The formulation of one-gene–one-enzyme hypothesis by Beadle and Tatum had sparked a revolution in biology.

The study of microorganisms has been central to the effort to understand life. We are still searching for answers to numerous fundamental problems. Where will the answers come from? As in the past, in large part from the study of microorganisms.

Investigation and inquiry

How is it then that in spite of an impeccable record of microbiology, its flag is at half-mast? Surprising as it seems, but professors and students genuinely interested in microbiology are indeed very few, as observed by Woese[2]. Is it that the contributions of microbes to biology is consciously or unconsciously covered up in the teaching of biology? Or, is it that the repeated assertion that 'what is true for *E. coli* is also true for elephants' generated the impression of the futility of gaining fundamentally new information, particularly when encyclopaedic information is already available on *E. coli*[18]?

The major fault in the teaching of microbiology (and other sciences) is that emphasis is placed on facts, instead of on observations and experiments. Invariably, the eminent microbiologists record that their curiosity and interest in microbiology were aroused from the excitement in collecting samples of some pond water or marine sediments or soil from some tropical forests or from some frigid place, from the suspense in waiting for microorganism to appear in an enrichment medium, from the fascination of the diversity of their types, from the microscopic observations of motile bacteria or of the intricate configuration of fungal spore walls, from the challenge of designing conditions for mass culturing a tractable organism, or from puzzling over the special features of the biochemistry which allows a particular microorganism to live in its ecological niche and cope with the environmental stress[22–25]. The eminence of these microbiologists comes from their realization that each microbial species is a unique, intrinsically valuable and a source of new knowledge.

Microbiological research

Extremophiles. Given the widely different level of scientific and technological development, the limited funds for science in our country, can we contribute to the vitalization of microbiology? The answer is a clear yes. I propose that research on extremophiles – the microorganisms which are indigenous to unusual habitats – may be a starting point. Perhaps, a fascination for the extremophiles as source of good biological questions and as new materials for new products and processes may usher a renaissance in microbiology in the country. India, with its diversity of natural habitats (desert, alpine and tropical soils, hot springs, porous rocks, glaciers, lakes, sediments of streams, caves, alkaline soils, brackish marshes, coastal lagoons, ocean, etc.) and a variety of man-created unusual habitats (composts, municipal waste dumps, sewage digestors, stagnant waters,

oil-contaminated areas, areas receiving industrial waste discharge, effluents of acid mines, corroded materials, etc.), presents great opportunities for the investigations of extremophiles. Because unusual habitats occur across the length and breadth of the country, research on extremophiles may well be one of the best ways to reach out and encourage scientists in the remote and less endowed institutions. Moreover, extremophile research is an area in which, because of the paucity of information, these scientists can contribute to the generation of knowledge, accomplish something and not feel frustrated. The pioneering studies of Thomas Brock on thermophilic microorganisms were done in remote places with simple facilities[26]. Could any one have predicted that the bacterium *Thermus aquaticus*, isolated in 1969 by Brock and Freeze from a hot spring in the Yellowstone National Park, USA, would in less than 20 years be exploited as a source of thermostable DNA-copying enzyme (*Taq* polymerase) and make the PCR technology an enormously successful technique? This enzyme and many restriction enzymes discovered from screening of thousands of bacteria and archaea are now churning millions of dollars annually in the developed countries. Witness the *Taq* polymerase controversy[27]. There was a time when people fought for the possession of women or for a piece of land. Now people are fighting for the rights to a bacterial gene product!

An understanding of life cannot be complete without a knowledge of the variations of biochemical design that evolved over many millions of years. Microbes offer unique experimental material for understanding the diversity of adjustments that they make to get adapted to their environment. Let us hope that the traditional areas of microbiology concerned with the study of the diversity of microbes and of their biochemical activities will be pursued vigorously.

1. Woese, C. R., *Proc. Natl. Acad. Sci. USA*, 1994, **91**, 1601–1603.
2. Woese, C. R., *Microbiol. Rev.*, 1994, **58**, 1–9.
3. Stanier, R. Y., Ingraham, J. L., Wheelis, M. L. and Painter, P. R., *The Microbial World*, Prentice-Hall, Englewood Cliffs, NJ, 1986.
4. Kluyver, A. J. and van Niel, C. B., *Microbes's Contributions to Biology*, Harvard University Press, Cambridge, 1954.
5. Fruton, J. S., *Molecules and Life: Historical Essays on the Interplay of Chemistry and Biology*, Wiley-Interscience, New York, 1972.
6. Brock, T. D. and Schlegel, H. G., in *Autotrophic Bacteria* (eds Schlegel, H. G. and Bowien, B.), Springer, Berlin, 1989, pp. 1–15.
7. Jannasch, H. W., in *Autotrophic Bacteria* (eds Schlegel, H. G. and Bowien, B.), Springer, Berlin, 1989, pp. 147–166.
8. Tunnicliffe, V., *Am. Sci.*, 1992, **80**, 336–349.
9. Stetter, K. O., Fiala, G., Huber, G., Huber, R. and Segerer, A., *FEMS Microbiol. Rev.*, 1990, **75**, 117–124.
10. Koch, R., Spreinat, A., Lemke, K. and Antranikian, G., *Arch. Microbiol.*, 1991, **155**, 572–578.
11. Gold, T. H., *Proc. Natl. Acad. Sci. USA*, 1992, **89**, 6045–6049.
12. Stetter, K. O., Huber, R., Blöchl, E., Kurr, M., Eden, R. D., Fielder, M., Cash, H. and Vance, I., *Nature*, 1993, **365**, 743–745.
13. Szewzyk, U., Szewzyk, R. and Stentröm, T.-A., *Proc. Natl. Acad. Sci. USA*, 1994, **91**, 1810–1813.
14. Pace, N. R., *Cell*, 1991, **65**, 531–533.
15. Barns, S., Fundayg, R. E., Jeffries, M. W., and Pace, N. R., *Proc. Natl. Acad. Sci. USA*, 1994, **91**, 1609–1613.
16. Bull, A. T., Goodfellow, M. and Slater, J. H., *Annu. Rev. Microbiol.*, 1992, **46**, 219–252.
17. Postgate, J. R., *The Outer Reaches of Life*, Cambridge University Press, Cambridge, 1994.
18. Perkins, D. D., in *More Gene Manipulations in Fungi* (eds Bennett, J. W. and Lasure, L. L.), Academic Press, New York, 1991, pp. 3–26.
19. Lehninger, A. L., Nelson, D. L. and Cox, M. M., *Principles of Biochemistry*, Worth Publishers, New York, 1993.
20. Cairns, J., Overbaugh, J. and Miller, S., *Nature*, 1988, **335**, 142–145.
21. Alberts, B., Bray, D. Lewis, J., Raff, M., Roberts, K. and Watson, J. D., *Molecular Biology of the Cell*, Garland, New York, 1994.
22. van Niel, C. B., *Annu. Rev. Microbiol.*, 1967, **21**, 1–30.
23. Stanier, R. Y., *Annu. Rev. Microbiol.*, 1984, **34**, 1–48.
24. Wolfe, R. S., *Annu. Rev. Microbiol.*, 1991, **45**, 1–35.
25. Pfenning, N., *Annu. Rev. Microbiol.*, 1993, **47**, 1–29.
26. Brock, T. D., *Thermophilic Microorganisms and Life at High Temperatures*, Springer, Berlin, 1978.
27. Lehrman, S. and Dickson, D., *Nature*, 1995, **375**, 378.

Ramesh Maheshwari is in the Department of Biochemistry, Indian Institute of Science, Bangalore 560 012, India.

Finding *patterns* in nature's maze: An endless quest

K. Chandrashekara

I am writing in response to 'Are patterns a rule in nature?'[1] by K. N. Ganeshaiah (KNG). Without doubt, reading it was quite a stimulating experience. However, I feel compelled to respond to a few issues raised in the paper; as the saying goes, it is better to debate an issue and not settle it rather than settle an issue without debating it.

Patterns: a question of scale

Are patterns a rule in nature? I find the question rather incomplete and bordering on the philosophical. What sort of pattern does the author have in mind? At what level of organization? At what temporal and spatial scales? To a physicist, chemist or for that matter any scientist, the answer to the question will be an assertive 'yes', given the fact that all matter in nature,

animate or inanimate, diamonds or DNA, represents a highly ordered arrangement of atoms. A historian sees a pattern in the evolution of civilization, an economist in share market fluctuations and a Coorg planter in coffee prices in the London market! They may, however, fail to see any pattern in short-range local events such as changing consumer tastes over a period of a few weeks or the variation in the onset of summer showers in April. Clearly, it is important to specify exactly what sort of pattern and scale one is dealing with before asking if patterns are ubiquitous.

What is a pattern: toward a definition

What is a pattern? As pointed out in the article, it is difficult to define what a pattern is. It is further argued that there may be disagreement among scientists on the signal component or specific feature of the system that is considered to reflect the pattern[1]. For this reason, the author says, it is difficult to define a pattern and that this constitutes a major difficulty in pattern finding in science. Perhaps not. Pattern is 'an arrangement or order discernible in objects, actions, ideas or situations' (*New Shorter Oxford English Dictionary*). Many of my colleagues when asked what the term pattern meant to them opined that the term pattern could be used to describe any phenomenon which is predictable. Unlike argued by KNG, it hardly matters what parameters are chosen to identify a pattern, because scientists from varied fields do recognize patterns in each other's fields though they may not fully understand the approach adopted in identifying the pattern.

In my opinion, what we probably have to split hairs about are the scientific merits and the criteria in classifying patterns as 'yes' or 'no' patterns. Taking the help of information theory, we can perhaps classify patterns. Many systems, be they organization of ideas or objects, carry information which has predictive value. Systems with '0' information content can be labelled 'no' patterns and those with information content of '1' as 'yes' patterns. Clearly, the information content will also influence the degree of predictability. Therefore, as pointed out by KNG, the patterns need not be discrete and there could be a continuum, at least in principle.

Patterns: good vs poor?

It is pointed out by KNG that there is a general agreement among scientists that patterns are an all or none phenomenon, notwithstanding his earlier argument that it is difficult to define what a pattern is and that it poses a major problem in pattern finding in science. He then expresses surprise that science immediately recognizes and appreciates '... demonstration of *absolute* "yes" patterns, more easily than that of "poor" or "no" patterns (italics mine)'. That is setting up a straw-man and shooting him down! As pointed out earlier, the crux of the matter is predictability and it is this feature which enables scientists to label something as a pattern. The reason why scientists do not appreciate a 'poor' pattern is because it is not explicitly labelled so. If a phenomenon concerning subatomic particles or stars has an information content which is greater than '0' (however small) and, therefore, has a certain level of predictability, it is labelled as a pattern.

Scientists are always aware of the inadequacy of a 'model' that explains patterns and it is unfair to say that no distinction is made between good, poor and no patterns. It is also true that as our knowledge about a phenomenon increases, the information content of the phenomenon may be revealed in greater detail and the predictability may also improve. This would result in a phenomenon being explicitly labelled as poor with hindsight, but for a given amount of knowledge and at any given moment there is merely a pattern or no pattern. Conversely, with increasing knowledge a 'yes' pattern may turn out to be a 'no' pattern. As an extreme example we can remind ourselves of Ptolemy's geocentric view of the universe, which was replaced by the heliocentric view – the well-known Copernican revolution.

A search for the 'ultimate particle' led Dalton to formulate the notion of indestructible atom; this he is known to have derived from the laws of visible heavenly bodies. But Humphry Davy was not convinced; he argued that 'there is no reason to suppose that any real indestructible principle has yet been discovered'. It took 50 years for science to prove Humphry right. However, by then, the concept of indestructible atom was used and elaborated by Gay-Lussac, Avogadro and Mendeleyev. The blow to indestructibility of atom came from two sources, the study of light and the discovery of electricity, the latter being totally unexpected[2]. The early history of physics and chemistry is peppered with many similar events that tell us three things; (i) that patterns when discovered are questioned; (ii) that even if the pattern is 'poor', considerable progress can be made in understanding nature and (iii) that the poor pattern may turn out to be a good pattern as we acquire more knowledge.

It is, therefore, not quite correct to say that science does not distinguish between good and bad patterns. When subjects in a psychological test are given a choice between various visual patterns and asked to sift the good from the bad, it is understandable, because they have a choice of 'alternatives'. Now how am I to say that the return of the Haley's comet once every 75 years, or the life-cycle of 13-year cicada is a good pattern or a bad pattern? If a good pattern is 'invariant' or has fewer or no 'alternative states', it again boils down to predictability or regularity, fewer the alternatives or alternative states, greater is the predictability. Which means we can now define a pattern in terms of predictability. If the predictability is greater than 0, we can claim to have discerned a pattern, however poor it may be. Once that is done, the scientists would like to take the predictability as close as possible to 1. This represents the range of patterns exemplified by Ganeshaiah's ink blot to a square or a circle. I argue that science has not instinctively picked the good patterns but has only so far described what is obvious; squares and circles – one of the two ends of the continuum. It would take more effort and time for science to traverse the entire range from both ends. Towards this the scientists are justified in using all the tools that are at their disposal and such attempts need not be dubbed 'frustrating search for patterns'.

Patterns and perception

KNG's major criticism is that our ability to detect patterns is limited by our perceptual abilities. This may indeed be so as philosophers right from Plato have emphasized. Given this handicap, how are we ever going to decide what is a 'good' pattern? A martian with better perceptual ability may throw out all our

'good' patterns as 'no' patterns! Our visual sense of a stick held in water deceives us to believe that the stick is bent at the water surface. How can we demonstrate that the stick is straight? Simple, by running our hand down the stick. Herein lies our folly and by extension the general tentativeness of all our discoveries. Because in our eagerness to prove an error in our sense of vision, we have trusted our sense of touch. In other words, how can 'deceptive tools' prove their own deception? However much we dislike it, we have to contend with our limited array of senses, suspect or not, to go about the job of discovering patterns in nature. Though we cannot assert that patterns abound everywhere in nature, we have to tap them using our limited tools. It is likely that as we polish nuggets of information and discover more and more of the hidden treasures, we will carve even more patterns from nature. If all this is the construct of the language and grammar of science, so be it. Roger Boscovich, a mathematician and physicist of the eighteenth century has gone beyond this and argued[3] that exploiting determinism to obtain complete knowledge is not possible because '... our human intellect is not equal to this task ...'.

It has been rightly argued that randomness is in the eye of the beholder and is a function of the power of our models to analyse the phenomena. Most of us have the experience of fitting our observed data points into one or more models. In the process, if we have been successful, it only shows that we have been able to explain the observed phenomena in some predictable way. I see no harm in such an exercise. Data fitting to wrench a pattern is a routine exercise through accepted rules of the game of science. Transformation of data is a legitimate practice. If log-transformation enables me to gather a greater comprehension of my observations (more patterned or less patterned), what is the harm? But for noise filters in radio and television, we would not be enjoying music, a pattern of sound.

Sure, as KNG points out, some amount of data or chips of information may be lost as noise in the process, but is it important? For instance, if a certain method of short-range forecasting of the possibility of rain in a particular week of the year is going to be predicted at some assured level of probability, even if it is not very close to 1, I am willing to lose any amount of information as noise through all the 'frustrating' jugglery!.

It required the mental acumen of great men to unravel patterns in nature about us where none was realized earlier. As Alexander Pope wrote, 'Nature and nature's law hid by night; God said, let Newton be and all was light'. To see woods from forest requires some skill and time. And seen it would be, definitely. So, if a hundred Newtons with their powers of senses identify patterns all around us, why complain. Certainly, it is our naivety to chide them that they are squeezing nature dry of 'no' patterns. Even in arts we need artists who see appealing patterns around them. I would welcome a hundred Michelangelos or Jakanacharis, who imposed *their* patterns onto stone – without their patterns the world would have been such a dull place.

Patterns: stochastic or deterministic?

I do not think that it is a general attitude of science to attribute an order-generating process to any pattern[1]. Science does not underplay or has disregard for the role of randomness in generating order. Stochastic processes are invoked as often, if not more, as are deterministic processes in describing patterns. Indeed, before the notion of chaos became an established concept, scientists addressed a complicated process as a statistical problem – a random process[3]. It is true that it took a long time for science to shed the blinkers of deterministic clockwork of the Newtonian world view – 'laws that never shall be broken'[3]. We now know that in a large number of physical systems, from heart beat to weather, the slightest uncertainty in the knowledge of the system at one moment will result in complete loss of information about the state of the system in a very short time and where randomness appears to prevail.

It is true that only recently have we begun to realize the power of randomness more explicitly in generating order, especially in biological sciences. Even here, for example, the so-called order we see in the evolution of organic diversity is a result of random or chance process. Evolutionary biologists agree that if the tape of evolution were to be rewound and run again, we cannot be certain that we will see an amoeba-like organism again; leave alone *Homo sapiens* or a *Tyrannosaurus rex*! Natural selection is not an order-generating process, it merely represents the success or failure of individuals of a species to respond to randomly changing local conditions.

The ability to find patterns has had a long evolutionary history. It was perhaps the ability of our australopithecine ancestors to find patterns in the movement of animals which helped them hunt, escape predators, and avoid other groups of competing individuals; later on, the same ability must have even enabled them to light fires. The success of australopithecines was due perhaps more to this than their erect posture or the ability to make tools! Human beings are known to make sense of their surroundings by making a 'mental representation' of what they see – an important feature of cognition. Given this it is not surprising that we try to find pattern in everything that impinges on our senses. I would even argue that this constant search for patterns contributed in no small measure toward the enlargement of the forebrain to which *Homo sapiens*, as a species, owes its success today.

I have deliberately not attempted to respond to the philosophical aspects of the question raised by KNG: simply because when I see a brick or a chair, I see *it*. I find it difficult to subscribe to the kind of philosophy which argues that a brick or a chair does not exist – it is only an interpretation of the human mind. It has to be. Even in Ganeshaiah's imaginary planet 'Binearth' we *need* a biologist from earth to study and interpret life on the planet. It is difficult to comprehend how the universe can be visualized without the existence of cognitive beings (for whom and by whom anyway?). The world is, what the mind sees. There is, in my opinion, no question of imposing anything.

Lastly, I do not think that science looks for patterns for its own survival. A very interesting pattern relating to the sounds produced by humans has been discovered recently. It concerns the case of two individuals with mild brain damage who were otherwise normal. It seems these two individuals were unable to pronounce vowels while being perfectly capable of pronouncing consonants. This provides yet another example of the 'pattern' of modular construction of human mind and

also suggests that categorization of sounds into vowels and consonants is a pattern. It may be *imposed* by human mind, so what? It could not be otherwise! Now if such pattern finding is going to ensure the 'survival' of neurologists and the growth of neuroscience, the 'last frontier in biology' we should only be too happy.

With surprises such as the one mentioned above being sprung on us from time to time, it may not yet be time to suspect that patterns are not a rule in nature. As Einstein said, 'the eternal mystery of the world is its comprehensibility', and toward this we should continue our search for patterns.

I am not sure if I have 'not settled' the issue by 'debate', but as Popper said, 'I may be wrong and you may be right and, by an effort, we may get nearer to the truth'.

1. Ganeshaiah, K. N., *Curr. Sci.*, 1995, **68**, 680–683.
2. Boorstin, D., *The Discoverers: A History of Man's Search to Know His World and Himself*, Vintage, London, 1985, p. 745.
3. Barrow, G. D., *Theories of Everything: The Quest for Ultimate Explanation*, Vintage, London, 1991, p. 223.

K. Chandrashekara is in the Department of Entomology, University of Agricultural Sciences, G.K.V.K., Bangalore 560 065, India.

MEETINGS/SYMPOSIA/SEMINARS

National Workshop on Biodiversity Conservation in Managed Forests and Protected Areas

Date: 29 November –1 December 1995
Place: Bhopal

Topics include: Biodiversity conservation in managed forests; Biodiversity in protected areas; Ethical and ecological considerations; Socio-economic issues; Institutional, policy and planning issues.

Contact: Prof. P. C. Kotwal
Indian Institute of Forest Management
P.O. Box No. 357, Nehru Nagar
Bhopal 462 003
Phone: 575716, 64268, 67851
Fax: 91-755-572878

Xth National Carbohydrate Conference

Date: 8–9 November 1995
Place: Vallabh Vidyanagar

Topics include: Plant and microbial polysaccharides; Isolation, characterization and their biological activity; Carbohydrate synthesis; Industrial polysaccharides; Production and utilization; Carbohydrate fermentation/Biotransformation.

Contact: Prof. M. N. Patel
Convener
Department of Chemistry
Sardar Patel University
Vallabh Vidyanagar, Kaira 388 120
Gujarat
Phone: (02692) 30411
Fax: 91-02692-46475

National Seminar on Trends in Chemical Education and Research

Date: 16–18 November 1995
Place: Madras

Topics include: Innovation in creative chemical education; Recent advances in certain areas of analytical and general chemistry; Current trends and progress in biological, industrial and environmental chemistry

Contact: Dr A. Srinivasan
Director
Protein Research Unit
Dept. of Chemistry, Loyola College
Madras 600 034
Phone: (O) 8276894 (R) 8270021

INSAT Master Control Facility: An insight

S. Rangarajan

An overview of the activities at the INSAT Master Control Facility (MCF) located at Hassan is presented. After a brief account of the INSAT services, the major subsystems of the spacecraft and their interrelationships are brought out. Important operations in different phases of the mission are described. The maintenance of these satellites in the geostationary orbit calls for periodic station keeping manœuvres. Facilities exist at MCF for carrying out these operations and also to provide support for the satellite users. A bird's eye view of the facilities is also included in the article.

The Master Control Facility (MCF) is the only mission control centre for the INSAT satellites – past, present and future. It is located about 8 km north of Hassan town on way to the historic monuments of Halebeedu. What immediately catches the eye for a newcomer is the number of antennae looking like giant-sized cups propping up over a green line of trees. Unlike the sunflowers in the fields around, these antennae at MCF seem to be pointed for ever into the vast emptiness of space. Actually they are looking towards and communicating with those satellites which have come to revolutionize our life-styles in the last decade. INSAT satellites are placed in the geostationary orbit, a circular one in the equatorial plane, at an altitude of around 36,000 km, a height at which the period of revolution synchronizes with that of earth's rotation. From there about 42% of earth's surface is accessible for line-of-sight communication using an antenna pointed in a fixed direction.

The services provided by INSAT satellites have grown steadily both in terms of quantity and in terms of new services. In addition to traditional applications in the areas of telecommunication, TV and radio broadcasting and imaging for meteorological purposes, INSAT has several unconventional services like the collection of local meteorological data from remote unattended platforms, relay of zone-specific disaster warning messages, standard time and frequency distribution, tele-education, search and rescue system, etc. In fact, some of the services are not widely known, for example, the high-quality sky radio channels which are being distributed as an add-on to the regional language TV transmissions.

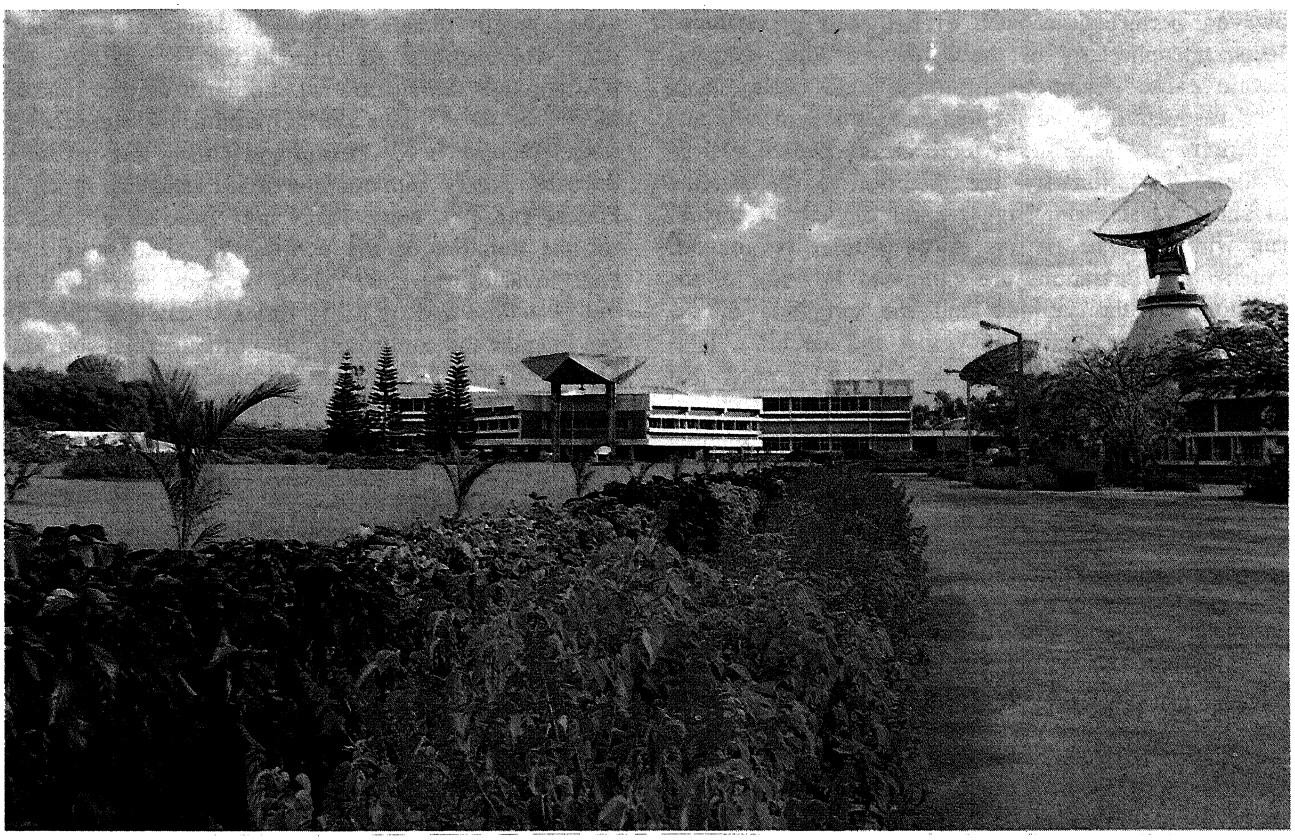
Spacecraft subsystems

To understand how these multidimensional services are provided and what MCF's role is in controlling the satellite, it is essential to have a broad understanding of the satellite's subsystems and the orbit. A satellite can be visualized as being composed of its payload and the mainframe or a platform to operate the payload, comprising many interlinked subsystems. The payload is that part of the satellite which specifically caters to the application. For example, the communication payload on INSAT provides the necessary functional capabilities to receive a radio signal from a ground station in a specified frequency band, amplify and retransmit back to ground in a different frequency band. There are a number of subsystems which are essential for any operational spacecraft. Electric power subsystem is generally composed of solar arrays for generating power, rechargeable chemical batteries for storing power and power control electronics for load regulation. The attitude and orbit control system (AOCS) is composed of a set of attitude sensors to sense the current orientation of the body, processing electronics and set of actuators to generate control torques to orient the payload in the desired direction. The thermal subsystem provides the required thermal environment for the different components/subsystems. Active control using heaters is usually available. The propulsion subsystem provides the capability to carry out orbit transfers, orbit maintenance, and provides for attitude control torques. The subsystem directly interfacing with the mission control centre is the telemetry, tracking and command (TT&C) subsystem. The communication is in specific RF frequencies in the GHz range. The telemetry subsystem centralizes the encoding of information from all the onboard subsystems and transmits them as a telemetry signal, probably the ever-running commentary about the on-board activities. Command subsystem provides the capability for receiving, decoding and distributing commands from the ground. Turn-around ranging provides the tracking data for orbit determination.

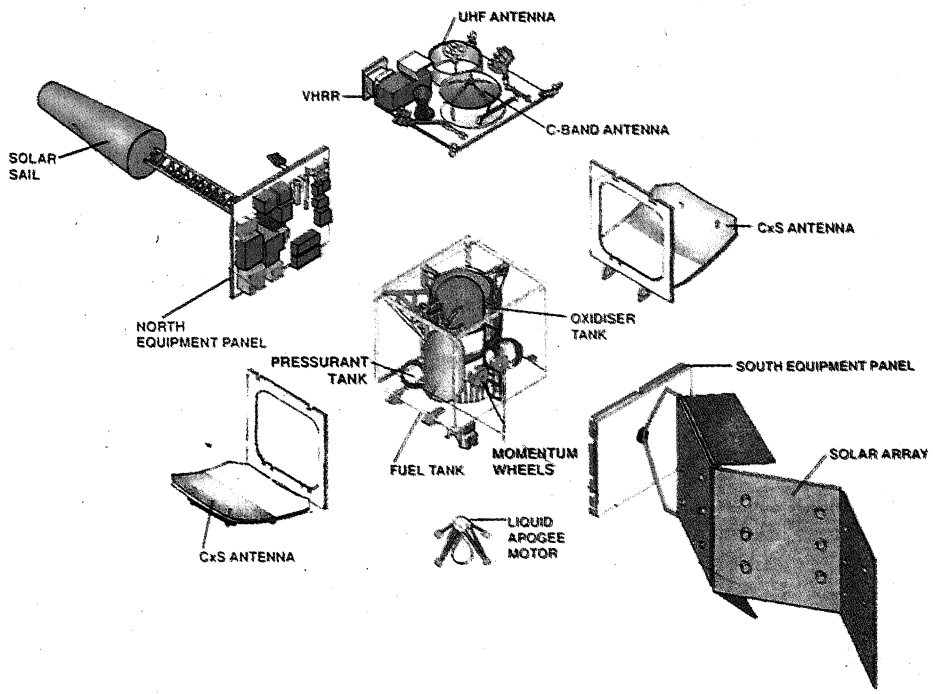
Orbit maintenance

A geostationary satellite has a designated longitudinal position in the equatorial plane over which it is placed.

S. Rangarajan is at INSAT Master Control Facility, Department of Space, Hassan 573 201, India.



A view of the Master Control Facility, Hassan.



INSAT-2B disassembled view.

However, perturbations such as luni-solar gravitational effects, nonsphericity of the earth and solar radiation pressure cause deviations from the ideal orbit. These affect, respectively, the inclination, semimajor axis and eccentricity of the orbit. If there is any departure from the ideal geostationarity, the direction to the satellite will show a variation. Bigger antennae have a smaller beam width and hence this effect is more pronounced in them. The deviations are expressed in terms of the excursions in the latitude and the longitude of the subsatellite point (vertically below the satellite). It is MCF's responsibility to maintain the orbit such that these deviations are less than 0.1°.

For this purpose the orbit is determined by sending a set of sinusoidal tones from MCF to the satellite, modulating it on a suitable uplink carrier. The tones are turned around by the satellite on a downlink carrier. By measuring the phase difference between the transmitted and the received tones, the distance to the satellite is calculated. This, coupled with the direction information as determined from the antenna pointing angles, provides the instantaneous position in three dimensions. Similar measurements carried out several times a day are sufficient for reconstructing the orbit. From the determined orbit, it is possible to work out the optimal date and time of orbit correction manœuvres. The inclination of the orbit, which affects the latitudinal excursions, must be limited to 0.1° in case of INSAT. This calls for an out-of-plane thruster firing once in about 80 days. The propellant consumed in each of these manœuvres is about 4 kg for INSAT, which has an on-orbit mass in the range of 1000 kg. Once there is insufficient fuel for this correction, the inclination of the orbit will build up, marking the end of mission life. The control of semi-major axis and eccentricity calls for in-plane thruster firing and the fuel requirements are an order of magnitude less. However, the frequency of corrections can be as high as once in two weeks. These manœuvres are critical activities carried out with the participation of subsystem specialists on a real-time basis. A typical manœuvre operation lasts about 4 h and the user services remain unaffected.

Orbital events

Precise orbit determination of satellite's orbit has other uses for preventive measures. If an observer views the celestial bodies from the satellite, he would naturally see them going around the satellite in various orbits, sometimes one occulting the other. Some of the orbit-related events are earth shadow, lunar shadow, sun/moon intrusion into the field of view of sensors, etc. As the evolution of the orbit is predictable to a reasonable accuracy, many of these events can be predicted well in advance and suitable actions planned. Some of the events greatly dictate the overall design of the satellite itself. For example, satellites in the geostationary orbit pass through the shadow of the earth around midnight local time for 45 days centred around the equinox. The longest eclipse is for 72 min and during this period the onboard battery supports the loads. Depending on the battery capacity, some loads may have to be shut off for the duration of the eclipse. The battery is to be charged using the solar power once the satellite comes out of eclipse and be ready for next day's eclipse. One can imagine the constraints put on the thermal subsystem of the satellite to maintain an equilibrium temperature for the critical subsystems in the absence of solar input for about an hour and to operate any heater, if at all, with tight power budget.

Link between control centre and satellite

It may be true that the design methods go on improving constantly to keep the satellite going, but still space environment is not yet completely understood to make the systems error-free. For the payload as well as the other subsystems of the satellite, it is essential to monitor the proper functioning. If ever a fault is detected, immediate correction is called for not only from the point of view of user services, but also for preventing the fault from worsening and propagating to other subsystems. For this purpose, it is essential to have knowledge of the current status of the satellite. A number of parameters pertaining to the different subsystems such as the temperature at different locations, operating status of equipment, pressure of fuel tanks, voltage, current and power consumed of selected units, angular information with respect to solar and other bodies, etc., are measured onboard and transmitted to the control centre. Some of these parameters are analog which are converted to digital data by sampling and quantizing. The sampling rates are different for different parameters. For example, temperatures which are generally slowly varying are sampled less frequently compared to attitude information, which is sampled more often.

The different parameters are multiplexed to form a composite telemetry stream. INSAT-2 satellite has about 400 analog parameters and 400 status parameters to be monitored. These are organized into a telemetry masterframe which repeats every 8 s. The slowly varying parameters occur only once in a masterframe. Each masterframe has 8 frames and each frame has 128 eight-bit words. Including the overheads required to identify the start of frame and the subframe within the masterframe, the telemetry has a bit rate of 1 kbps. As the transmission from the satellite is at a very high frequency, the telemetry data are first modulated on a subcarrier which, in turn, modulates the RF beacon. INSAT uses PSK modulation on a 32 kHz subcarrier

and PM modulation on a 4 GHz carrier. On ground, the entire sequence has to be undone, namely, receive the beacon, recover the subcarrier, obtain the bit stream, synchronize the frame, identify the words and bits, convert to engineering units, and display the data. Besides being useful for mission control, the telemetry data also provide important feedback to the designers.

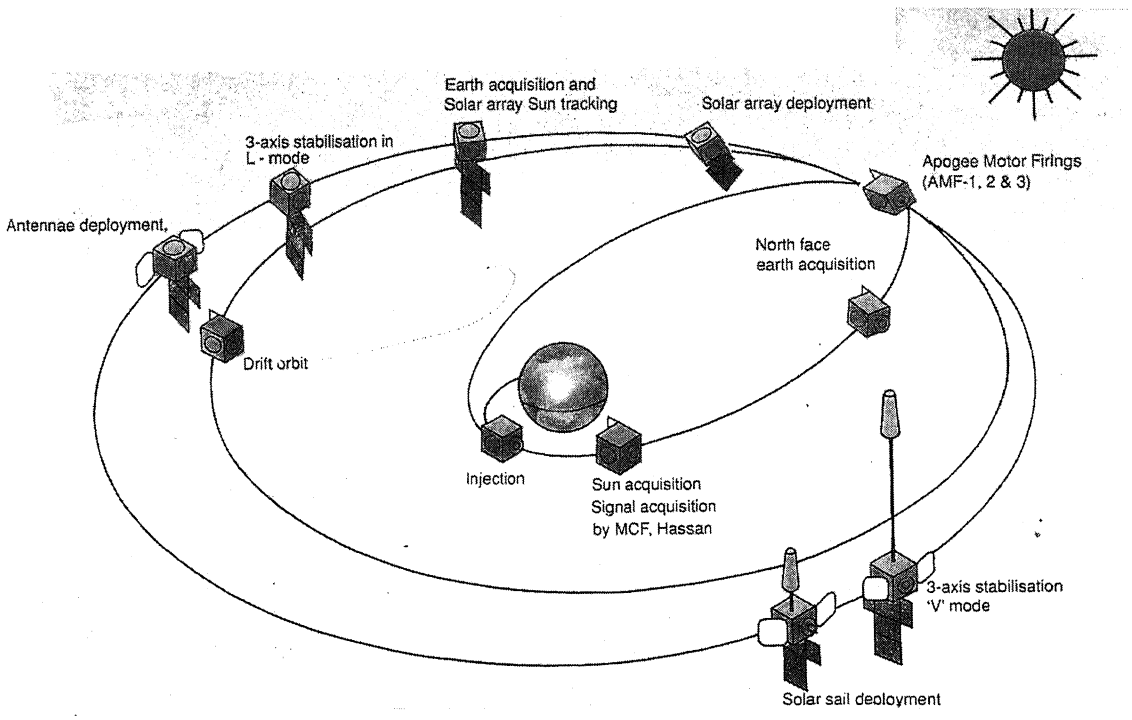
As the number of parameters is large, continuous surveillance of the data is carried out by the mission computer, which can present the data in graphical and tabular forms and also alert the operator in case of any abnormality. Even one bit of data being abnormal will be a nerve-racking experience to the operating crew.

The telecommand system provides the capability for receiving, decoding and distributing commands sent from the control centre. Commands are required for operating payloads, switching on/off heaters, firing thrusters, selecting a redundant system, etc. Here again the command code is a set of bits which are to be transmitted on a RF carrier. INSAT uses a 100 bps data with FSK/FM modulation on a 6 GHz carrier. Unlike telemetry, it is important to guard against even a single bit error in the transmission. For this purpose forward error control coding is used.

Launch phase operations

In today's technologies, satellites are not just thrown in space. They are designed and built on ground but transported to the destination in space. Before reaching the final slot, each satellite, like INSAT, goes through a launch phase wherein many critical activities are undertaken. The launch and early orbit phase is full of excitement and MCF bubbling with activities. In the first one or two weeks, the major operations include firing of big rocket engines for orbit raising, opening out of antennae and solar panels, an elaborate testing of payloads, etc. In INSAT the transfer to the geostationary orbit is carried out by using a liquid apogee motor which gives a thrust of 450 N. This motor has to be fired in three parts for about 100 min in all, resulting in a propellant consumption of about 1000 kg. This must be compared with the propellant requirement for the next ten years of on-orbit operation, which will be around 200 kg. Deployment of the appendages is also critical as any failure will result in a total loss of mission.

So far MCF has supported six launches and in each case the signal from the satellite was received minutes after the satellite was put into orbit by the launch vehicle. This orbit is a highly eccentric one with perigee as close as 200 km from the ground and apogee at the geosynchronous height. Consequently, there is a large motion of the satellite relative to the ground calling for tracking antennae. Also, a given earth station cannot have continuous radio contact in successive orbits. For this purpose, additional tracking stations are deployed and a communication interface established with the tracking network. The launcher agency is also linked by dedicated communication channels so as to monitor



Mission sequence.

carefully the prelaunch operations which go on for 7–8 weeks. Prior to the launch the entire hardware at MCF undergoes revalidation. The mission control software is tested extensively and the operations crew are trained to carry out the sequence of events (SoE) pertaining to the mission. The simulation and rehearsals are conducted in several stages and last a few months.

Normal phase operations

At the end of launch phase operations a satellite and its payloads are tested for their functioning and declared operational from its slot in space. Then onwards, the mission life of the satellite begins and the operations undertaken in this period are called normal phase operations. Spacecraft operations call for considerable planning and critical analysis so as to ensure flawless execution of commands and a fast response in the event of any contingency, pinpointing the problem area in case of a malfunction. Based on the user requirements nominal operations are detailed as an event schedule on a time line. Unforeseen behaviour is taken care of by issuing operations directives generated by the subsystem engineers. Equal importance is given for trend analysis and performance monitoring to detect even minor anomalies at the earliest. Emphasis is also laid on generating periodical reports as well as anomaly reports.

Though the operations may look quite routine for a casual observer during the normal on-orbit phase, the stakes involved are so high that it is important to monitor continuously the state of health of the satellite. A careless command or unwatchful operator can bring the end of life to a mission. The parameters pertaining to attitude control are rather important in this context. The attitude and orbit control system (AOCS) of INSAT in the normal on-orbit phase is designed around the concept of biased angular momentum. By spinning two flywheels an angular momentum of about 130 N m s is maintained, giving gyroscopic stiffness to the body. The disturbance torques are taken care of by adjusting the wheel speeds, making use of the principle of conservation of angular momentum. Unidirectional disturbances may result in the wheel speeds going to extreme values and hence at such times thrusters are fired (typically for 16 ms) to bring the wheel speed within the operating range. Even though the attitude control system works based on a closed-loop onboard control, it is important to monitor the system parameters from ground. Even a small error of 1° in the pointing will result in a shift of antenna beam by about 600 km on the surface of the earth, affecting all the user services. Such misorientations can arise for a variety of reasons like an anomalous thruster firing, a single event upset in the processor code, attitude sensor failure, etc. If any of these were to occur, quick corrective actions are called for. Depending on how high the body rates are, the recovery process may take anywhere from 15 min to 10 h. At such times of loss of attitude lock, the experts assemble at the control centre within 15 min irrespective of the time of occurrence of the anomaly.

MCF facilities

MCF is an integrated control facility with multiple earth stations co-located with mission control centre. It presently has three full-motion antennae with trans-receive capability in C-band (6/4 GHz). There are a number of limited-motion antennae to support payload test in C-band, extended C-band, S-band (2 GHz) and UHF (400 MHz). It also has a boresite facility for zero range calibration and RF testing. The interface between earth stations and satellite control centre (SCC) is at 70 MHz. SCC houses the base-band equipment for telemetry, tracking and command, data processing system, timing systems, recording facilities and communication systems. The computer systems include PDP 11/70s and VAX systems. In keeping with the present-day trend, the future missions will be supported on distributed processing system. Besides the workstations, telemetry is being processed on PCs with suitable add-on cards developed in-house. The mission control software can be broadly divided into two categories – one for real-time data acquisition, processing, health monitoring and trend analysis, and the other for processing the tracking data, determining the orbit, predicting the orbital events and planning correction manœuvres.

A centralized power system provides uninterrupted power supply for each functional area. The captive generating system consists of six generators of 310 kVA capacity each. A variety of air-conditioning systems provide the required environmental control. It is a matter of great pride that all these systems are maintained in-house.

Initially it was a necessity to grow a garden around the facility to avoid collection of dust all around. Over the years the garden has been developed so well that it won a prize in the horticultural competition at the State level.

The facility started with just about 40 persons in the year 1980. Now even the security staff belonging to the CISF exceed this number! The present staff strength of MCF is around 300, drawn from different engineering disciplines, technical branches and administrative areas. At any time 18 persons are on round-the-clock shift duty headed by an Operations Manager. He or she will coordinate with the users on real time as required. The spacecraft controllers undergo a well-structured training programme for six weeks followed by on-the-job training for a similar duration. It is noteworthy that MCF was

asked to organize a course for such controllers working for Arabsat, a multinational organization controlling similar satellites. If MCF has been able to achieve smooth and flawless operations for all these years, it is only due to the personal involvement of everyone of the staff members who takes pride in his or her job with an excellent team spirit.

MCF has been responsible for several innovations such as the survival of INSAT-1B through the eclipse season in spite of the loss of both the onboard batteries, use of attitude control thruster firings for orbit adjustment, detailed on-orbit measurement of antenna footprint, etc. MCF has also taken the lead in indigenization programmes as well as technology transfer to the industry.

Into the future

INSAT system is still in the stage of growth in terms of introduction of newer payloads and newer services. Correspondingly, the augmentation of the facilities at MCF will go on for the next few years. For example, presently, new earth stations are being built for testing Ku-band (11/14 GHz) communication payload and Mobile Satellite Service (MSS) payload. In the next three years a launch of INSAT is planned every year. There will also be experimental missions carried on the test flights of GSLV. Each launch calls for calibrating, checking and revalidating the performance of every subsystem to ensure fail-safe operation. The launch seasons are so hectic that the MCF personnel have very little time for other activities. However, the routine operations tend to be less glamorous, but provide useful feedback to the designers. ISRO is one of the few organizations in which the satellite design and operations are carried out under the same umbrella. This is sure to pay rich dividends in the long term.

Vegetarianism: A nutritional appraisal

K. T. Achaya

Vegetarianism in India is perhaps 3000 years old. What is the extent of flesh consumption at present? How do vegetarian diets stand up to scrutiny in the light of modern nutritional knowledge? What of the future? These are the questions addressed in this paper.

NOWHERE in the world except in India is the concept of vegetarianism really commonplace, or accepted without explanation. Nor is this surprising, since the vegetarian ethos was propounded here by the Vedic Indians as early as about 1000 BC. There are many today who are a product of total vegetarian eating for as long as hundred generations, and survive, reproduce and live full and satisfying lives, both physically and mentally. So it has been done, and can be done. There is room, however, to review the current trends in the use of animal foods, and the practice of vegetarianism, in relation to the knowledge of scientific nutrition that has been unravelled over the last century or so.

Proportions and causes

Census figures of 1971 record the proportions of vegetarians that exist in the various states of India. Among the high-vegetarian-population states (expressed as percentages of the population) are Gujarat 69, Rajasthan 60, Punjab–Haryana 54 and Uttar Pradesh 50. At medium levels come Madhya Pradesh 45, Karnataka 34, Maharashtra 30 and Bihar 24, while the low-vegetarian-population states are Tamil Nadu 21, Andhra Pradesh 16, Assam 15, and Kerala, Orissa and West Bengal 6 each. The overall figure for vegetarianism in the whole country in 1971 was 28%.

What is the nature of this vegetarianism? Reliable dietary data have been collected for the years 1979–81 by the National Nutrition Monitoring Bureau, Hyderabad, in 10 states spread all over India[1]. In the rural areas of these 10 states, the consumption figures for fish and other flesh foods (including eggs)[1] are shown in Table 1. Everywhere, fish intake is much higher than that of meat, the averages for each of these for the 10 states being about 10 g of fish and just 3 g of meat a day. Since vegetarians form about one-quarter of the popula-

K. T. Achaya resides at Flat GB, Wood Creek, 115, 3rd Main Road, Defence Colony, Bangalore 560 038, India.

Table 1. Nonvegetarian food intakes: Rural dwellers (g/consumption unit/day)

Region	Fish	Other flesh foods	Total
North			
Uttar Pradesh	3	2	5
Madhya Pradesh	0	0	0
South			
Kerala	37	1	38
Tamil Nadu	3	8	11
Karnataka	1	2	3
Andhra Pradesh	15	7	22
East			
West Bengal	17	4	21
Orissa	12	5	17
West			
Maharashtra	11	3	14
Gujarat	1	0.5	1.5
Mean, about	10	3	13
Weighted mean for nonvegetarians, about	13	4	17

tion, the daily averages for nonvegetarians as a whole are 13 g of fish and 4 g of meat, a total of 17 g per head per day. In other words, in India, the nonvegetarian item by and large represents just one more vegetable, only a small part of the total food. Larger amounts of fish are eaten in coastal states like Kerala (37 g), and West Bengal, Andhra Pradesh and Orissa (17–12 g)[1]. In Goa and the Andaman Islands, data from other surveys conducted around the same time show that per capita fish consumption is as high as 100 g a day, with practically no meat at all[2].

Two factors are involved in vegetarianism. One is the ethical or religious belief, to which have recently been added economic and environmental concerns. The other is income, because meat and, slightly less so, fish are expensive items of food. Table 2 shows intakes of nonvegetarian foods for various income groups in Kerala, Tamil Nadu and Gujarat[1]. In Kerala, 94% of the population are nonvegetarians: fish eating is widely prevalent, and income is not a very major determinant. In Tamil Nadu, religious belief and the income effect are both depressants of nonvegetarian intake, but the reason seems to vary at different levels. Upper income groups are vegetarians through belief, lower ones through income scarcity. More meat than fish is eaten at all levels, a distinct and somewhat surprising reversal of the general rule for a coastal state. Gujarat too has a long sea coast, but the large populations of Jains and Vaishnavites are responsible for very low consumption of both fish and meat at all income levels[1].

Simply the fact of living in cities seems to influence the nature of food intakes in several respects[3]. The dietary data in Table 3, showing the consumption of nonvegetarian food in Trivandrum, Madras and Ahmedabad, the capitals of the three states that have just been discussed, illustrate in exaggerated form the rural deductions just made[3]. In Trivandrum, high-income groups take both fish and meat, other groups mainly fish. In Madras, again the distinct preference among upper-income groups for meat over fish is apparent even when both can be afforded; is not, fish is consumed. Quantities are low in Madras upper-income groups because of ethical vegetarianism. In Ahmedabad, there is far more nonvegetarian food consumed than the average 1.5 g/day in rural Gujarat (Table 1). Again, meat is preferred over fish, reflecting the choice of the cosmopolitan (and perhaps not always Gujarati) upper-income groups of that city.

Table 2. Nonvegetarian intakes: Rural dwellers (g/consumption unit/day)

Region (population group)	Fish	Other flesh foods	Total
Kerala			
High-income	45	9	54
Middle-income	47	1	48
Low-income	33	0	33
Tamil Nadu			
High-income	3	8	11
Middle-income	4	9	13
Low-income	1	7	8
Gujarat			
High-income	0	0	0
Middle-income	2	0	2
Low-income	1	0	1

Table 3. Nonvegetarian intakes: Urban dwellers (g/consumption unit/day)

Region (population group)	Fish	Other flesh foods	Total
Trivandrum			
High-income	36	30	66
Middle-income	47	3	50
Low-income	41	0	41
Madras			
High-income	4	15	19
Middle-income	14	13	27
Low-income	17	8	25
Ahmedabad			
High-income	4	19	23
Middle-income	1	5	6
Low-income	2	10	12

The complexions of regional diets

So far we have focused on the nonvegetarian part of Indian diets. Lets us now look at the rest of the diets, shown in Table 4 for the rural areas of four states situated in the north (Uttar Pradesh), south (Tamil Nadu), east (West Bengal) and west (Gujarat) of the country[1]. In all these diets, nonvegetarian items form only a very small part: rural Indian food everywhere in the country

is essentially vegetarian. The quantities of cereals and pulses consumed in all states are fairly similar, though the cereal used will differ, and to a lesser degree the pulse. Uttar Pradesh would have mostly wheat, Tamil Nadu and West Bengal mostly rice, and Gujarat rice, wheat and bajra. Intakes of roots and tubers, and of vegetables, are low in Tamil Nadu and Gujarat, and high in West Bengal. Consumption of milk and its products is low in West Bengal and exceptionally high in Gujarat. The levels of calories, proteins and fats calculated from the above dietary components[1] are shown at the end of the table, but these are better considered a little later, along with the quantities of other minor nutrients derived from these diets.

These nutrient intakes[1] are shown in Table 5, along with the 1989 recommended daily intakes for each nutrient laid down by the Indian Council of Medical Research[4]. In the new norms, the weights of an Indian reference man and woman have been set higher (at 60 and 50 kg, respectively) than the current averages in order that the levels of nutrients recommended become targets to be reached by the better-built citizens of tomorrow. This may make some of the actuals appear more deficient than they really are for the smaller people. Generally speaking, calories, proteins and fats are fairly well met even in average Indian rural diets. Calcium and iron are apparently adequate, though iron represents a special case, which will be discussed later. Vitamin A intakes are really low, and are reasonably satisfactory only in Bengal, thanks to the high consumption of green leafy vegetables, a cheap and excellent source of this vitamin. Other nutrient requirements are more or less met. It would appear that, generally speaking, Indian-style cereal-based diets, though containing very little nonvegetarian foods, are nutritionally fairly satisfactory, and can be made even more so with greater buying power and some dietary knowledge and alteration. Let us now look into our foods and cooking practices in greater detail.

Table 4. Average regional rural diets (g/consumption unit/day)

	Uttar Pradesh	Tamil Nadu	West Bengal	Gujarat
Cereals	497	518	572	438
Pulses	47	35	23	32
Roots/tubers	79	39	93	46
Vegetables	140	106	235	96
Milk and its products	75	82	39	199
Fats/oils	4	11	8	21
Sugar/Jaggery	7	17	11	33
Animal foods	5	11	21	2
Calculated nutrients				
Calories	2115	2196	2580	2333
Protein, g	70	54	63	67
Fat, g	25	36	32	51

Indian food practices

Many traditional Indian habits have come down to us from very ancient times, and are based on the food materials actually available, on deductive observations in the past of their effects on health, and on the overall Indian food ethos. These practices have much to

Table 5. Nutrients present in Indian rural diets

	Uttar Pradesh	Tamil Nadu	West Bengal	Gujarat	ICMR recommendation
Calories	2115	2196	2580	2333	2700
Proteins, g	70	54	63	67	60
Fat, visible plus invisible, g	25	36	32	51	20[d]
Calcium, mg	426	609	493	546	400
Iron, mg	29	26	33	25	28 (women 30)
Vitamin A, mcg	207	211	495	264	600 (lactation 950)
Vitamin C, mg	41	39	91	35	40 (lactation 80)
B vitamins					
Thiamine, mg	2.1	0.9	1.1	1.9	1.35
Riboflavin, mg	1.2	0.8	0.8	1.2	1.62
Nicotinic acid, mg	21.6	12.1	17.9	14.1	17.60
Pyridoxine, mg	—	—	—	—	2.0[a]
Folate, mcg	—	—	—	—	100[b] (pregnancy 400)
B_{12}	—	—	—	—	1.0[c] (lactation 1.5)

[a]Wheat and bajra: rich; rice: poor.
[b]Liver and eggs: rich; cereals, green vegetables, banana, orange: moderate.
[c]Mostly from animal foods and bacterial activity (see text).
[d]Visible fat only.

commend them. Everywhere in India cereals and pulses are eaten together: rice with sambhar, chapati with dal, and combination cooked foods like the idli, dosai, dhokla, pesarattu, khichdi and holige. There is an important protein complementation effect involved[4,5]. Cereal proteins carry good levels of the two essential amino acids cystine and methionine, and pulse proteins that of the amino acid lysine. When both are taken together, these amino acids complement each other, raising the total protein eaten to one of higher quality. If percentage marks are given, eggs, meat and fish proteins would score about 80 and milk about 72. A cereal–dal combination would score about 65, though each protein in itself would score only 50–55. Add to this some milk or curd, and there is not much to choose in protein quality between vegetarian and nonvegetarian diets taken as a whole. A second point about Indian-style diets is that they are rich in dietary fibre[5]. This is a recent term which means all complex carbohydrates that are not broken up by the digestive enzymes and are, therefore, not absorbed, thus passing into the stools. Today, many diseases of the lower colon, like cancer and constipation, are being linked with an insufficiency of dietary fibre, and hence of long intervals between evacuations. Indian diets rich in cereals and vegetables do well on this count. So do they in terms of fat[6]. Rice and wheat carry about 2–4% of invisible fat each, and consumption of the usual 500 g of cereal a day would give 10–20 g of fat right away. Indian diets have been computed to contain a total of as much as 25 g of such invisible fat a day on an average[6]. Fermented foods have much to commend them. The digestibility of both carbohydrates and proteins in foods like idli and dhokla is very high, since these entities have been broken down during fermentation[7,8]. Any iron that is present becomes ionized and, therefore, available to the body. Curds carry microflora developed during fermentation, which serve to replenish those in the lower intestine and are helpful in digestion and in warding off infection.

Today the virtues of liquid unsaturated vegetable oils are being publicized in the western countries. India has had a long tradition of the use of such oils in daily cooking. Ghee has had a place of veneration that current nutritional science does not tend to endorse[9].

Raw foods are valued because they are one way of ensuring a good supply of a nutrient like vitamin C, which is quickly destroyed by air, light and heat. Western salads of raw vegetables and fruits are meant for this purpose, and in India we have long cherished our raw chutnies, pacchadis, raitas and kocchumbers made up of ground green ingredients. In fact, cooking without fire was a distinct Vedic category of food preparation[10–12] and included items like pickles of lime and mango with spices and salt, or sugar-preserved items in which nutrients are well conserved for year-round use. Sprouting of mung, chana and ragi is a well-established Indian practice, and often the products are eaten raw, which further conserves vitamins. Vitamin C and several B vitamins are nearly doubled on sprouting and iron is released from bound forms[7,8,13]. An age-old practice thus finds modern nutritional vindication.

Indian historical injunctions regarding hygiene have a thoroughly contemporary connotation. One was obliged to wash hands and the mouth before and after eating, eat from disposable plates to avoid cross-infection, and not to sip water from a glass but pour it in. Leftover food was not to be eaten, which in pre-refrigerator days made sound sense. There were two kinds of foods in the Aryan ethos[12]. *Kaccha* foods were everyday foods, eaten daily, like rice and dal; these boiled foods were to be cooked and eaten only within the kitchen by the family members. *Pucca* foods could be taken outside, and also shared; these were usually fried foods, which, as we know today, have a much longer keeping quality, and are not prone to infection. Sociological theorizing was in this instance a basis for sound nutritional practice.

Animal foods in nutrient terms

Proteins of high quality were once considered the major virtue of animal foods. There is truth in this in terms of amino acid quality, but as we have seen, blending of proteins from cereals and pulses compensates fairy adequately. The quantity of protein in fresh meat and fish is 18%, whereas most dals carry between 20 and 23%. However, because of swelling and bulkiness, it is easier to consume say 100 g of meat than to eat the same weight of dal, and the gas-forming nature of most dals is also a drawback[8,13].

There is a bit of a mystery about iron. It has repeatedly been shown that from nonvegetarian diets, there is high absorption of any iron present in the food, whereas the same quantity of iron from a cereal-based diet is absorbed only to the extent[14] of about 3%. Also, from the same diet, a pregnant woman will absorb about 5% of the iron and a woman suckling a child[4] 7%. Some sort of feedback mechanism operates, and vegetarian diets do face a problem of low absorption. As a result, most Indians are slightly anaemic by western standards, and consumption of iron-fortified salt is being recommended even on a national basis. Studies in large groups have shown excellent and sustained improvement in blood iron levels after taking such salt[15].

Vitamin B_{12} is an 'animal' vitamin, present in high amounts in liver and to lesser extents in meat. From purely vegetarian diets it is difficult to get enough of this vitamin to meat normal needs, and yet there is little evidence of deficiency in India by way of pernicious anaemia[7]. This has been attributed to several causes:

the higher absorption (70%) of the vitamin from food by vegetarians than by meat-eaters (16%), the presence of bacteria (which synthesize vitamin B_{12}) on the surfaces of green vegetables and in water supplies carrying decaying vegetable matter, and the poor hygienic practices prevalent in India[4]. The quantity of the vitamin needed daily is very small, only one mcg a day. Yet many vegetarians from India who migrate to England frequently develop signs of vitamin B_{12} deficiency like pernicious anaemia; the type of bacteria populating their small intestines has been found to change, and the number of bacteria drops considerably, probably reducing B_{12} synthesis in the small intestine markedly[16]. Also, a large number of antagonistic factors may come into play, like the use of more processed foods, an excess of vitamin C and copper, and excessive medication. Cholesterol is another product of animal metabolism, found in eggs and meat, and also in milk, butter and ghee. Consumption of these is low in India, but the high incidence of cardiovascular disease among Indians living in England, once tentatively attributed to oxidized cholesterol found in heated ghee freely used by Indians there, is now being linked[9,17] to the *trans* fatty acids present in ghee to the extent of about 4%. Fears regarding cholesterol consumption have sharply depressed intakes of butter (which is also a saturated fat) in advanced countries, perhaps without justification. Cholesterol in milk itself is at a low level and milk is not considered a major source.

Global concerns

In India, in the recent years there has certainly been a trend towards a higher degree of nonvegetarianism among traditionally vegetarian families. Many factors are operative. Better purchasing power among the middle class is one of them, and the aping of the habits of the so-called upper class, the executive and the film star. To restless younger people, it represents part of a desire to break traditional shackles, especially when they travel abroad alone and undergo the travails of finding appetizing vegetarian food when on the move. Whatever it be, the phenomenal growth during the last two decades of the egg and chicken industry in India is a reflection of the spread of nonvegetarian eating habits among Indians.

The west, on the other hand, seems to be moving steadily towards vegetarianism. Both in Britain and the USA, some 7–10% of people, especially the younger generation, are believed to have become vegetarians. Many reasons are given for this. One is the humanitarian concern of *ahimsa*. A more important reason is the realization that the animal is an inefficient converter of feed materials into meat or milk. The same food grains which an animal is given could feed several human beings, and the production of a typical western non-vegetarian diet of say 3000 calories has actually used up many times that value in terms of feed calories. There is a growing realization of the fact that our earth is a single, finite spaceship in which the whole of humanity will sink or swim together: food must be found for everyone, and animal food is too energy-expensive to meet the need.

What of the future? One can only hazard a guess that a new genetic revolution in food plants is on its way, which may increase grain yields to an extent that can hardly be visualized. If grain production from the same agricultural land area goes up dramatically, and if the population of the world gets stabilized, the economic argument will lose strength. Whether the ethical or moral one will continue to prevail remains an open question.

1. Report for the year 1979, Section II, National Nutrition Monitoring Bureau, Hyderabad, 1979.
2. Statewise dietary surveys in India in the 1980s, Food and Nutrition Board, New Delhi, 1989.
3. Report on the urban population, National Nutrition Monitoring Bureau, Hyderabad, 1982.
4. Nutrient requirements and recommended daily allowances for Indians, Indian Council of Medical Research, New Delhi, 1989, pp. 43, 68, 97.
5. Gopalan, C., Rama Sastri, B. V. and Balasubramaniam, S. C., revised and updated by Narasinga Rao, B. S., Deosthale, Y. G. and Pant, K. C., *Nutritive Value of Indian Foods*, National Institute of Nutrition, Indian Council of Medical Research, Hyderabad, 1989 (revised edn), pp. 3, 8, 23, 25.
6. Achaya, K. T., *J. Sci. Ind. Res.*, 1987, **46**, 112; 1995, **54**, 91.
7. Rajalakshmi, R., *Applied Nutrition*, Oxford and IBH, New Delhi, 1974, 3rd edn., pp. 107–109, 236, 266.
8. Deosthale, Y. G., *Interfaces between Agriculture, Nutrition and Food Science* (ed. Achaya, K .T.), The United Nations University, Tokyo, 1984, p. 171.
9. McKeigue, P. M., Miller, G. J. and Marmot, M. G., *J. Clin. Epidem.*, 1989, **42**, 597.
10. Khare, R. S., *The Hindu Hearth and Home*, Vikas, India., 1976.
11. Khare, R. S., *Culture and Reality*, Indian Institute of Advanced Study, Simla, 1976.
12. Achaya, K. T., *Indian Food: A Historical Companion*, Oxford University Press, New Delhi, 1994, pp. 62–65.
13. Udayasekhara Rao, P. and Belavady, B., *J. Agric Food. Chem.*, 1978, **26**, 316.
14. Narasinga Rao, B. S., Vijayasarathi, C. and Prabhavati, T., *Indian J. Med. Res.*, 1983, **77**, 648.
15. Narasinga Rao, B. S., *Nutr. Found. India Bull.*, 1981 (April), 1.
16. Ballentine, R., *Transition to Vegetarianism*, The Himalayan International Institute of Yoga, Science and Philosophy of USA, Honesdale. Penn., 1987, p. 169.
17. James, W. P. T., *Nutr. Found. India Bull.*, 1995, **15**, (April, No. 2), 1.

Cursive English script recognition – A knowledge-based system

P. V. S. Rao and T. M. Ajitha

Script recognition by computers becomes extremely difficult due to variability at the input arising from a variety of reasons. Most existing recognizers are based on statistical pattern recognition techniques and require extensive training. This system uses domain knowledge to extract the broad shape features while filtering out the deformations and other variations in shape as 'noise'. This eliminates the need for any type of training. The performance of the system (even in a writer-independent task) is in the range of 95–100%, which compares very favourably with the current state of the art. This approach can be used for Indian language scripts too.

SCRIPT recognition – recognition of handwritten words and characters by computers – is receiving a lot of attention nowadays due to its practical applications in many areas: office automation, bank cheque processing, postal address and ZIP code recognition, signature verification, and, in general, document and text recognition. It is, in addition, a challenging area of research.

Script recognition can be accomplished using either of the following two approaches: recognizing words as integral units or visualizing each word as being composed of discrete characters. This distinction between word- and character-based recognition systems is particularly significant for cursive English script, where each word is not merely a string of discrete characters but a more or less continuous curve where character shapes may get modified to some extent. Recognition of words in terms of individual characters becomes difficult because decomposing a word into character length segments is a nontrivial problem. Segmentation into characters is not needed for a word level recognition system; on the other hand, the number of classes (each word will be a distinct class) increases very substantially in this case.

The different methodologies used for script recognition (character- or word-based) include:

- Point-by-point global comparison (comparison of all pixels in the image).
- Global transformations involving Fourier transformations, moment calculations and rotation according to the principal axis of inertia.
- Extraction of local properties such as lines, end points, T-junctions, corners and strokes.
- Analysis in terms of curvature which involves curve following, detection of concavities and geometrical analysis.
- Structural methods (decomposition of each character into its constituent elements, topological description and reduction of each character into a graph).

P. V. S. Rao and T. M. Ajitha are in the Computer Systems & Communications Group, Tata Institute of Fundamental Research, Homi Bhabha Road, Colaba, Bombay 400 005, India.

Overview

How can a computer-based recognition system 'learn', i.e. acquire knowledge about the characters and script of a language? How does it organize and utilize this acquired knowledge for actual recognition? The task of handwritten character and word recognition gets complicated due to variability arising from context effects, sloppy writing, subject-to-subject differences in writing style, etc. In fact, even the same person writes quite differently at different times. It is for such reasons that even the best trained optical readers – the human eyes – make some 4% mistakes when reading handwritten script in the absence of context information. According to Mantas[1]: 'Handwritten character recognition is not a simple task as it may appear. Errors in reading hand prints are caused by the infinite variations of shape resulting from the writing habit, mood, style, education, region of origin, health, social environment and other conditions of the writer, as well as other factors such as writing instrument, writing surface, scanning methods and finally, of course, the machine's character recognition algorithms.' The complexity of the task increases progressively as we go from fixed-font recognition of printed text (recognition of specific fonts) to recognition of single unconnected handwritten characters and finally to connected cursive (handwritten) scripts.

The underlying presupposition behind all recognition approaches is that there exist certain invariant properties which define each pattern in the task domain in a unique way and that these properties (parameters) can be determined in a quantitative sense.

In a general case of pattern recognition, each sample or pattern can be represented as a point in an n-dimen-

sional parametric space where each dimension corresponds to one parameter (Figure 1). The set of parameters required for representing the individual patterns at the input would naturally depend on the complexity of the problem domain. Ideally, various instances of each class map into points which cluster into a compact neighbourhood in this space. Also, there would be no significant overlaps between regions defining the classes. Both these are necessary conditions for pattern recognition systems to perform well. These facilitate deriving a 'mean' representation for each input class, determined by the input instances (belonging to this class) encountered at the time of training.

For instance, in the case of the minimum-distance approach, the 'distance' between an input pattern and the 'mean' or 'class prototype' is a measure of the dissimilarity between them. The input pattern will be recognized as the class prototype 'nearest' to it. In the maximum-likelihood approach, the likelihood that a test sample may be a member of a cluster or class is determined on the basis of probabilistic distributions in the parametric space. The input pattern is put into the class with the maximum value for the probability of occurrence.

Many existing recognition systems use statistical pattern recognition techniques based on minimum distance, maximum likelihood, hidden Markov models[2], neural network models[3,4], etc. They are trained 'by exposure' to representative samples; the system tries to extract representative models for each of the script characters (or words) as encountered in real life. Understandably, these models would be distorted versions of the corresponding ideal character shapes.

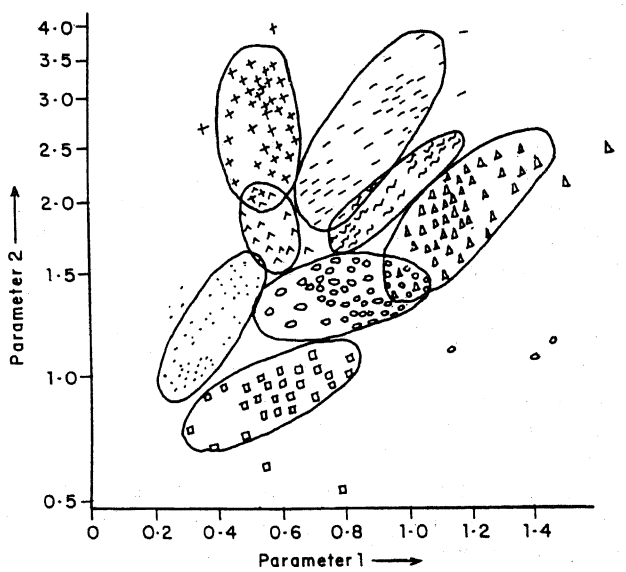


Figure 1. Clustering in parametric space.

Figure 2 shows some representative patterns 'learnt' by a neural network classifier proposed by Schomaker[4], for the characters */a/*, */b/*, */g/*, */k/*, */n/*, */o/*, */p/*, */r/*, */w/*, */x/*, */y/*, */z/*. It can be seen that the mean representation or class prototype is rather distorted and is quite different from the 'ideal shape' for each particular class (character). This is because while developing a 'statistical model' for the respective input classes, the classifier is, in fact, modelling not only the canonical shape behind it but also the variations (or 'noise') encountered at the input level.

In contrast to the statistical approaches mentioned above, our feature-based approach[5,6] is based on the fact that for each script character or word, there exists notionally an ideal pattern – the copy book version – which the writer is attempting to reproduce while writing. The variations or distortions that occur are viewed as added noise. Essentially, our system represents each individual character/word (input class) by the corresponding canonical shape (prototype). The task of the recognition system is thus to map a noisy real-life input (Figure 3 *a*) into the copy book counterpart (Figure 3 *e*). This is done in two steps.

Step 1: Filter out (local) shape-related distortions and retain only the broad (global) shape features of the input shapes by choosing an appropriate set of parameters which preserve only the very essential information regarding the shapes (Figure 3 *b*). It can be easily verified that there exists only one way of linking the feature points while preserving the order. The shape distortions in Figure 3 *a* have been eliminated in the reconstructed pattern (Figure 3 *c*).

Step 2: Filter out the distortions relating to size, shape and slant. Figure 3 *e* is such a reconstruction. (This is accomplished by quantization of parameter values as shown in Figure 3 *d* and is discussed in greater detail later.)

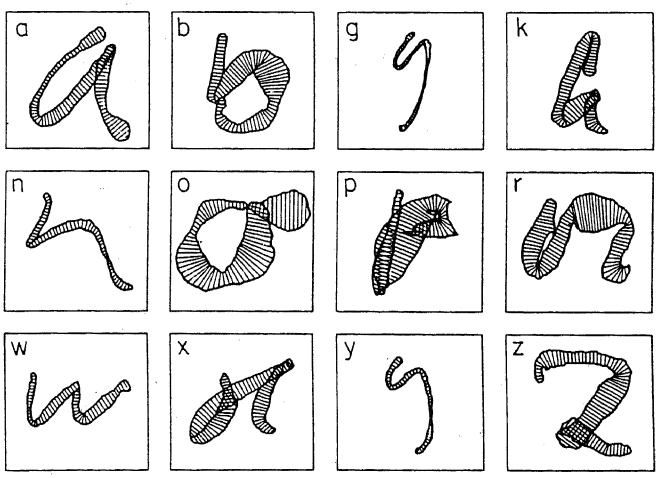


Figure 2. Abstract shapes derived by Kohonen net trained on samples written by a single subject (taken from ref. 4).

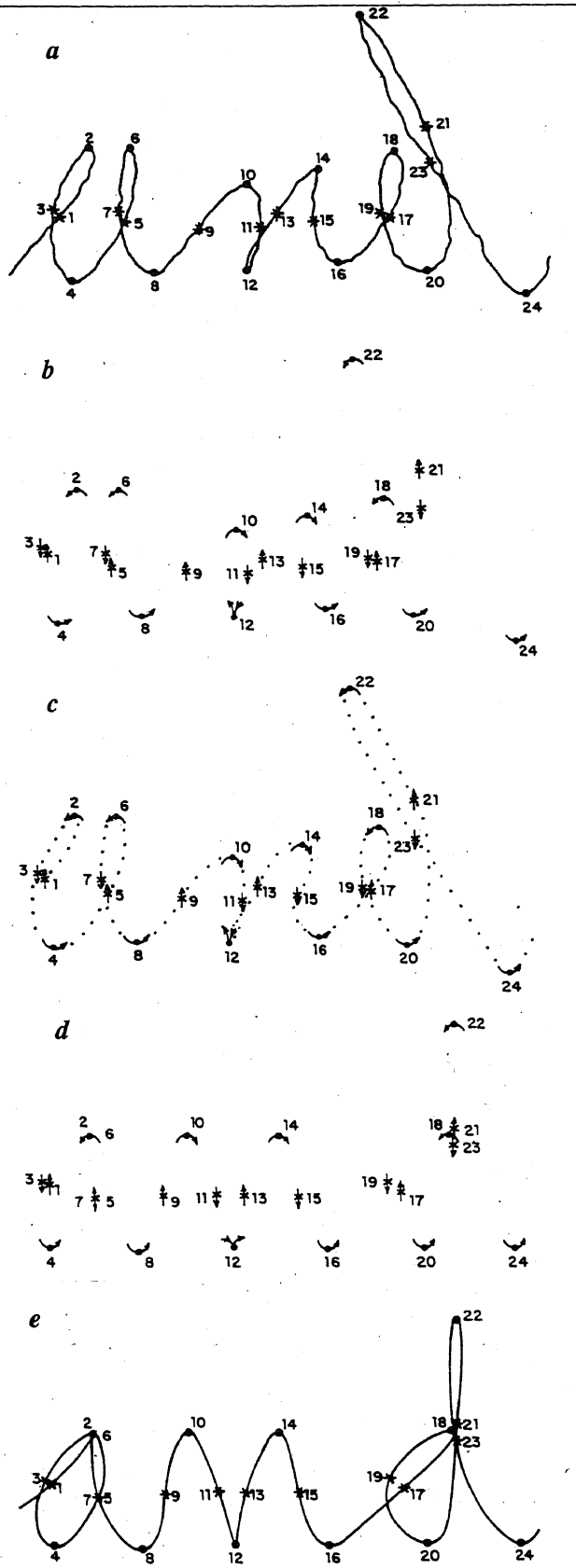


Figure 3. Feature extraction and reconstruction after normalization. *a*, Original pattern. *b*, Information captured by the feature set used. *c*, Reconstruction based on information in (*b*). *d*, Feature set after quantization. *e*, Reconstruction using quantized feature set.

Thus, this feature-based approach helps us to filter out the deformations, thereby ignoring the actual and detailed shapes of the character and to retain only broad shape features. Consequently, there is no need for analysing different samples corresponding to each individual character since each of them carries the same broad shape features. This makes training superfluous in this approach.

Selection of parameters

As mentioned earlier, recognition needs to be performed on the basis of the attributes of the patterns involved; these are represented as values of specific physical parameters of the signal. Selection of an appropriate minimum set of parameters is thus an important step for any script recognition system, and, in general, for all pattern recognition systems. For example, the selected set of parameters should not be sensitive to the style (and mood) of the writer (or to any of the other variations that occur in practice). It would be ideal if these parameter values remain invariant irrespective of contextual effects, variation in size, and placement of individual characters within a word. A character level recognition system should also be insensitive to all possible variations of individual character shapes when they are concatenated to form a word. This is a very important requirement for cursive script recognition in English.

In the current feature-based approach, the parameters selected permit the retention of broad shapes by capturing the archetypal characteristics of each individual character, and filtering out as noise all writer-specific, context-dependent and other distortions such as those relating to shape, size and slant. We accomplish this by heavy quantization of the parameter values (Figure 3).

Feature-based recognition system

In our parametric model, we use features concerned with points of minimum radius of curvature (minima) and maximum radius of curvature (maxima) – the feature points – for representing the characters. The feature vector representing an individual character consists of the points of minimum radius of curvature – the minima – of the character. The attributes we use are the location of the minima and the direction of curvature at these points.

Figure 4 shows the word */andy/* and the minima and maxima points; the minima bear even numbers and the maxima are the points with odd numbers.

The attributes of the minima and maxima are insensitive to local fluctuations in the shapes of the actual curves. A recognizer using these attributes would, there-

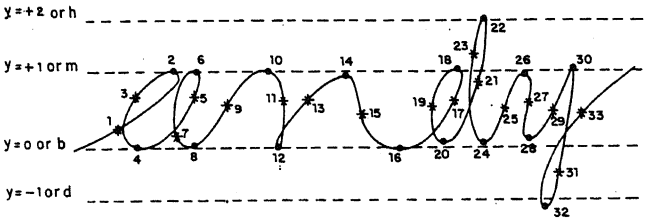


Figure 4. Feature points for the word */andy/*.

fore, be reasonably robust. It is, in principle, possible to train a statistical recognizer to recognize script words using this information alone. Though insensitive to shape fluctuations, such a recognizer would still be sensitive to the fluctuations in slant, size and position. Such a recognizer would, therefore, need a large number of training samples. To obviate this, it would be desirable to 'standardize' the character size which manifests itself as horizontal and vertical distances between successive minima. This standardization can be conveniently done by warping the characters to fit into a uniform grid, i.e. by quantizing the X and Y coordinates of the minima points. This takes care of variation in character sizes. This has, in fact, been done in Figure 3 *d*. Similarly, quantizing the slopes of the maxima would eliminate the fluctuations in slant angle while writing.

Each minimum on the curve is characterized by position (X and Y coordinates) and curvature. The curvature is quantized to three values to carry the direction information: 'a' for anticlockwise, 'c' for clockwise and 's' for sharp turns (where the curvature could be either clockwise or anticlockwise). The Y coordinate value is quantized to four possible values: 'b' and 'm' (baseline and midline) for characters such as */c/* and */n/*; 'h' and 'l' (highline and lowline) for characters with ascenders and descenders such as */h/* and */g/*. X coordinate values are expressed as relative displacements as follows:

- for minima at the baseline and midline, with respect to the nearest left minimum at the same level.
- for minima at the highline and lowline, with respect to nearest left minimum at any level (because it is unlikely that there will be an adjacent left minima at the same level in these cases).

This helps us to distinguish between */d/* and the combinations such as */el/* or */il/*. Figure 4 and Table 1 illustrate this scheme. Minima 2, 6 and 18 have the value 'a' (anticlockwise) for direction while minima 10, 14, 26 and 31 have the value 'c' (clockwise) for direction. Minima 12 and 30 have the value 's' (sharp) for direction. Minimum 12 has been traced in clockwise direction (as a loop). It could alternatively have been traced as a sharp turn as in the case of minimum 30. This would have made the curvature of minimum 12 anticlockwise; hence, the need for classifying it as

Table 1. Feature vectors for */a/* and */n/* in */andy/*

		character a					character n				
Max.	Serial No.	1	3	5	7	9	9	11	13	15	17
	Slope	U	D	U	D	U	U	D	U	D	U
Min.	X	f	f	n	f		f	f	f	f	
	Y	m	b	m	b		m	m	b	b	
	Curvature	a	a	a	a		c	s	c	c	
	Serial No.	2	4	6	8		10	12	14	16	

'sharp' rather than as clockwise or anticlockwise. X has the value 'n' for minimum 6 because it is very near to minimum 2 at the same Y level and at its left, and 'f' for minimum 10 since it is far from its left neighbour. The X value helps to distinguish between */a/* and */u/*, */g/* and */y/*, */o/* and */v/*, etc. The X coordinates are expressed not as absolute values but relative to each other, so that they remain unchanged even after concatenation of individual characters.

Maxima points on the curve are labelled by the slope angle quantized to the following values: up, down, right and left. The actual thresholds used for quantization are:

$0 \pm 20°$: right (R)
$180 \pm 20°$: left (L)
$90 \pm 70°$: up (U)
$270 \pm 70°$: down (D)

All the above parameters are extracted from single samples of individual letters. The sequence of values constitute the feature vector for the character.

The feature matrix for a word is obtained by concatenating the feature vectors for individual characters forming the word in the same sequence.

A lexicon consisting of feature matrices for all the words in the vocabulary is compiled from the feature vectors of individual characters. The test word written by the subject (which is to be recognized) is similarly analysed to yield the corresponding feature matrix. This is then compared with the feature matrices of all words in the lexicon. Recognition is accomplished on a best match basis, i.e. either the exact match or, failing this, the nearest match.

Details of implementation

Data acquisition and preprocessing

Both reference and test data are collected using a graphics tablet connected to the serial port of an IBM-PC-compatible system and feature vectors are extracted

therefrom. Specimen samples for the 26 English lowercase letters are collected from a single subject. Feature vectors extracted from these letters constitute the reference data for the recognition system. The writer has to write the input pattern on the graphics tablet using the digitizer pen. The tablet digitizes the handwritten pattern and sends a stream of X and Y coordinates to the host computer through the serial port. These data are stored in the ASCII file format. To facilitate extraction of height information, it is necessary to ensure that the script is in a straight line and that the characters are of a reasonably standard size. To this end, the upper and lower reference lines are provided on the writing surface as in ruled copy books for children. The subject is required merely to write broadly within these lines. There is no need for anything like copy book precision.

Radius of curvature extrema – the feature points on the curve – can be located using one of several alternative methods. One of the easiest ways to locate minima (and maxima) is by using the fact that the stylus travels fast and almost in a straight line near the maxima but is slow and almost doubles back on itself as the curve turns nearly 180° around minima. The chord distance (not arc distance) between two sample points separated by constant time would, therefore, be minimum around minima and maximum in the neighbourhood of maxima. The relevant parameter values can be extracted for each minima and maxima and quantized easily.

Handwritten characters are usually not vertical in their orientation; they tilt to a greater or lesser extent, usually to the right. The tilt or slant remains more or less invariant for each person and is one of the features that characterize an individual's handwriting. It would seem very desirable to eliminate this slant so that we can deal with more standard – essentially vertical – characters. This can be done by means of a number of slant estimation and correction processes. It turns out that, except for extreme slants, this is not really essential, since our method of parameter representation and quantization for script characters is robust enough to tolerate even major slants in the writing.

Comparison and recognition

For recognizing the given input word, we narrow down the list of prospective candidates (dictionary words or reference patterns) on the basis of a rough match and perform a detailed matching only within this short list. The number of minima in a word is used as the criterion for this short-listing. All dictionary words where the mismatch in the number of minima exceeds 3 are rejected. If an exact match is found within the short list, recognition is successful. Failing this, the best match is identified from the short-listed set of words. We use a comparison procedure based on dynamic programming which permits efficient matching of feature vectors even in the presence of spurious minima and absence of genuine minima in the test pattern. The input word is recognized as the word in the short list which matches the test pattern best. Recognition fails if either the short list is empty or if, for all words in the short list, the mismatch during detailed matching turns out to be excessive.

Experimental results

As stated before, each character is represented by a feature vector, a sequence of labelled minima and maxima points. Each minimum is characterized by (a) location (quantized) and (b) direction of movement (clockwise, anticlockwise or sharp turn) while each maxima is labelled with the direction of movement. Recognition is successful if an exact match is found within the short list. Failing this, the best match is selected as the candidate.

Data sets were collected from 10 (male and female) subjects from different age groups and educational backgrounds. Each subject was asked to write the 63 words of the vocabulary (word length varied between 3 and 7 characters) twice on the graphics tablet. The vocabulary included words such as 'fill' and 'fell' which are very smilar to each other. The reference character set (26 lowercase alphabets) was taken from a different subject. Standard lexical entries for each word in the vocabulary were generated from these by appropriately concatenating the feature vectors of these individual letters. Recognition results (using this 'common' set of lexical entries as standard and the word sets written by each subject at the test set) are summarized in Table 2.

Each subject was also asked to write ten sets of the word */and/*. These were pooled together and used in a separate recognition experiment. In this case the overall recognition score was 93%. Figure 5 includes some of the samples which were recognized correctly despite the significant shape deviations in many of them. This demonstrates the robustness of the approach followed.

Table 2. System performance

	Recognition scores (%)	
Subjects	Word set 1	Word set 2
Subject 1	95	97
Subject 2	97	97
Subject 3	98	98
Subject 4	97	97
Subject 5	97	98
Subject 6	97	100
Subject 7	100	100
Subject 8	100	97
Subject 9	100	100
Subject 10	100	100

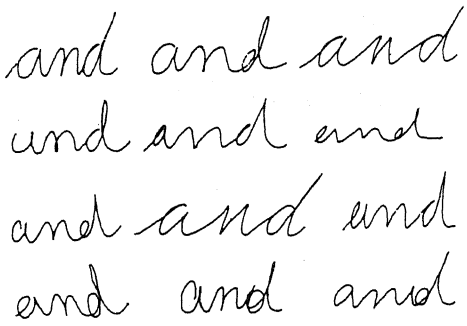


Figure 5. Samples of the word */and/*.

Discussion and conclusions

In this feature-based approach for script recognition, we introduce an appropriate empirical parametric model for cursive English script. The minimal set of parameters used permit the retention of the shapes of the script characters, while ignoring the context-dependent, writer-specific and other distortions in shape, size and slant of characters. Resynthesis based on extracted parameters was successful in recreating the shapes of connected words and individual characters. This demonstrates the validity of our approach.

This is a knowledge-based system; knowledge about the task domain is contained implicitly in the set of parameters chosen for representation and the manner in which they are processed. Needless to say, any recognition system needs knowledge regarding the task domain at some stage before it can start 'recognizing' the input. Statistical systems start without any *a priori* knowledge. Consequently, such classifiers acquire this knowledge through elaborate training. In the case of hidden Markov models and neural networks, the topology chosen implicitly incorporates some knowledge about the task domain into the system. This, however, is not enough to dispense with the need for training. In the current feature-based approach, we use knowledge about the task domain to choose an appropriate parametric model which extracts and retains only the canonical or prototype shape information for specimen words in the lexicon on the one hand and for the test words on the other hand; in both cases, the model filters out real-life shape deformations as 'noise'. This obviates the need for training of the type needed for statistical or neural-network-based approaches. Only an 'initialization' with one sample for each of the 26 lowercase script characters is required prior to recognition. The performance of the system does not degrade significantly even when the initialization and test samples are written by different subjects. The recognition scores are in the range of 95–100%. This demonstrates the robustness of the approach.

The present system is able to handle vocabulary sizes of the order of a few hundreds of words. Though the parameter set selected here is ideally suited for the Roman script, the underlying principle is general enough to adopt for script recognition in other languages, say Indian languages. The set of features or parameters used for representation might, of course, vary across languages. We are developing feature-based recognition systems for some Indian languages and preliminary results are encouraging. Machine recognition of Indian language scripts becomes a particularly attractive avenue for research due to the complexity of the task of adapting Indian scripts to the qwerty keyboard.

1. Mantas, J., *Pattern Recogn.*, 1986, **19**, 425–430.
2. Kundu, A., He, Y. and Bahl, P., *Pattern Recogn.*, 1989, **22**, 283–297.
3. Rao, P. V. S., Proceedings of International Conference on Automation, Robotics and Computer Vision (ICARCV '90), Sept. 18–21, Singapore, 1990, pp. 1237–1241.
4. Schomaker, L., *Pattern Recogn.*, 1993, **26**, 443–450.
5. Rao, P. V. S., Proceedings of IAPR Workshop on Machine Vision Applications, Nov. 28–30, Tokyo, 1990.
6. Rao, P. V. S., *Sadhana*, 1994, **19**, 257–270.

ACKNOWLEDGEMENTS. We thank Mr Thomas C. Jose, Mr Yashowardhan R. Sowale and other student trainees for their help at various stages of experimentation and to other colleagues who helped in data collection. We are also grateful to Mr R. S. Thakur for his neat drawings and timely help.

REVIEW ARTICLE

Wave–particle duality and Bohr's complementarity principle in quantum mechanics

D. Sen*, A. N. Basu and S. Sengupta

Department of Physics, Jadavpur University, Jadavpur, Calcutta 700 032, India
*Present address: Barasat Govt. College, Barasat, West Bengal, India

Interest in Bohr's complementarity principle has recently been revived particularly because of several thought experiments and some actually performed experiments to test the validity of mutual exclusiveness of wave and particle properties. A critical review of the situation is undertaken and it is pointed out that the problem with mutual exclusiveness arises because of some vagueness in the conventional formulation. An attempt is made to remove this vagueness by connecting the origin of mutual exclusiveness to some principles of quantum mechanics. Accordingly, it becomes obvious that to contradict complementarity principle without contradicting quantum mechanics would be impossible. Some of the recent experiments are critically analysed.

IN classical physics energy is transported either by particles or by waves, and particle and wave are the sole ultimate categories of physical entities. Particles are described by Newton's equation of motion and a wave by a space–time field function which satisfies a differential equation – the so-called wave equation (or field equation), and there is no mix-up. But the classical 'particle' is described in quantum mechanics (QM) by a field function (ψ) satisfying Schrödinger's equation which has a wave-equation-like flavour. Similarly, classical 'waves' are described by quantum field equations in QM which may be recast in a Schrödinger equation[1] for the wave function of the field quanta. The basic difference in the mathematical description of the two distinct classical entities disappears and we are in for the typical wave–particle duality in the world of quantum mechanics.

The wave–particle duality according to Feynman[2] 'contains the only mystery of the theory', while Bohm[3] considered it as one of the crucial 'new features of primary significance' introduced by the quantum theory. Eddington has coined the name 'wavicle' for microparticles and the term 'wave function' has been chosen to emphasize the wave–particle duality, considered generally so fundamental in quantum mechanics. Wave–particle duality is pointed out even in a quark model based on finite-dimensional quantum mechanics[4]. On the other hand, Popper[5] writes '... I nevertheless believe that the quantum theory is in a very definite sense a particle theory (here I disagree with Schrödinger) and in a sense which excludes a duality, analogy or complementarity between particles and waves'.

Notwithstanding such comments, the notion of duality in the physical description which grew in the years following Einstein's work in 1905 continues to spur fresh debate in the wake of some recent experiments on matter–wave interferometry and quantum optics.

Almost 20 years after the seminal work of Einstein, de Broglie extended the wave–particle dualism to material particle in 1924 and signalled the formulation of wave mechanics. Acknowledging the significance of the work, Einstein[6] immediately wrote back that 'de Broglie has lifted a corner of the great veil'. However, in the following years multiplicity of dualist interpretation led Bohr[7] to 'harmonize the different views, apparently so divergent' and to offer a general description of 'atomic phenomena' in the framework of the complementarity principle.

The recent glut of critical comments regarding the validity of Bohr's complementarity principle (BCP) or, more precisely, the conceptual content of the wave–particle duality in the light of new experiments, however, makes it clear that much of the controversy arises because of the prolixity of the formulation, which contains some amount of vagueness (see Box 1). In January 1949 Einstein[8] complained that 'despite much effort which I have expended on it, I have been unable to achieve the *sharp formulation* of Bohr's principle of complementarity' (italics ours). In our view two major shortcomings of the conventional formulation of BCP are: (1) There is no guideline to decide whether a suggested pair of variables/properties should be considered complementary. (2) No precise quantum-mechanical principle is adduced from which the incompatibility can be derived.

When Bohr remarked '... any given application of classical concepts precludes the simultaneous use of other classical concepts which in a different connection are equally necessary for the elucidation of the phenomena'[9], he kept one important question unanswered. If an experiment is performed to measure some properties and use classical concepts for the elucidation of the

phenomenon, how do we ascertain which concepts are to be used and which to be precluded? This difficulty arises mainly because Bohr formulated the complementarity principle without addressing the principles of QM. We can find a parallel in the principle of uncertainty. Heisenberg formulated it without using QM. But later it was demonstrated that the uncertainty principle is implicit in the formalism of QM. Some physicists seem to accept BCP as independent of QM, and its relation to QM remains illdefined. Any attempt to show that mutual exclusiveness (ME) in BCP can be violated without violating QM presumes the view that BCP is independent of QM. It is surprising that no serious attempt so far has been made to see if BCP can be formulated in a way so that it becomes completely consistent with the principles of QM.

In a recent communication[10] we made such an attempt and pointed out that both the uncertainty and the complementarity principles are implied within the formalism of QM. Also it must be clearly understood that they add nothing to the formalism. Instead, they bring out explicitly certain features of QM not quite obvious from the formalism itself. In addition, they focus our attention on some sharply contrasting situations in classical and quantum mechanics and thereby help us realize their fundamental differences. According to Scully *et al.*[11], 'complementarity distinguishes the world of quantum phenomena from the realm of classical physics'.

In the quantum-mechanical description of a physical entity which is classically described either as a particle or a wave, we find that pairs of variables/properties exist which are mutually exclusive in the sense that a *sharp simultaneous knowledge of both is impossible.* Actually these are termed as complementary pairs.

In the above statement we have made ME a *defining property* of complementary pairs. On going through the writings of Bohr on BCP, it is evidently clear that ME is intended to be the essence of complementarity and we conclude that if for a system mutually exclusive variables do not exist, the system is a purely classical

Box 1.

N. Bohr on complementarity principle

'An adequate tool for a complementary way of description is offered by the quantum-mechanical formalism... It must be remembered that even in the indeterminacy relation we are dealing with an implication of the formalism which defies unambiguous expression in words suited to describe classical physical pictures. Thus, a sentence like "we cannot know both the position and the momentum of an atomic object" raises at once questions as to the physical reality of two such attributes of the object which can be answered only by referring to the conditions for the unambiguous use of space–time concepts on the one hand and the dynamic conservation laws on the other hand. While the combination of these concepts into a single picture of a causal chain of events is the essence of classical mechanics, room for regularities beyond the grasp of such a description is just afforded by the circumstances that the study of the complementary phenomena demands mutually exclusive experimental arrangements.' – Bohr, N., in *Atomic Physics and Human Knowledge*, Science Editions, New York, 1961, p. 41.

Box 2.

Pros and cons

1. '... the development of modern physics has enriched our thinking by a new principle of fundamental importance, the idea of complementarity.' – Born, M., *Physics in my Generation*, Springer, New York, 1969, p. 171.
2. 'Complementarity denotes the logical relation, of quite a new type, between concepts which are mutually exclusive, and which, therefore, cannot be considered at the same time – that would lead to logical mistakes – but which nevertheless must both be used in order to give a complete description of the situation.' – Rosenfeld, N., *Nature*, 1961, **190**, 384–388.
3. 'I have, for a long time, adopted the idea of complementarity in the realm of quantum physics, whilst at the same time realizing that it was inadequate. In recent years, I have been led to regard the concept of complementarity with increasing suspicion.' – de Broglie, L., in *The Current Interpretation of Wave Mechanics*, Elsevier, Amsterdam, 1964, p. 7.
4. 'This illogical conflation of concepts was the great quantum muddle which led to the emergence of the notorious wave–particle duality and hence the complementarity interpretation.' – K. R. Popper's view on the Copenhagen interpretation as presented by Jammaer, M., in *The Philosophy of Quantum Mechanics*, Wiley, New York, 1974, p. 448.

one. In every complementary pair he has discussed, Bohr stressed the ME aspect. But a gap was created when he said that wave properties are complementary to particle properties – it is absolutely necessary to know in a given experimental set-up which particle properties are to be regarded as complementary to which wave properties. Classical waves and particles have many common properties. As Bacry[12] puts it, '. . . , there are situations where the two aspects compete to interpret a phenomenon'. But in this respect Bohr's specific statement is significant: 'It is evident, however, that any concrete wave picture is as unable to account for basic experience regarding the individuality of the electron as a corpuscular picture for the superposition properties of radiation fields'[13]. This statement very clearly indicates that the complementarity Bohr meant was between properties of superposed and unsuperposed states and, as we shall see presently, these properties definitely satisfy ME.

In ref. 10 we made a precise statement of BCP in the spirit of the statements made by Bohr and in this paper we intend to examine critically in the light of this precise statement the thought experiments or the actual experiments which claim to violate the Bohrian notion of ME.

That complementary pairs of variables exist for quantum physical entities, no one doubts, and when we measure the precise position of an electron, we know that the concept of momentum has to be excluded. This knowledge comes from the principles of QM. Similarly, if we are doing an interference experiment, say double slit it is the which-state information that is complementary to the appearance of the interference pattern. We want to apply the same criterion in the so-called wave–particle complementarity, identifying the quantum principle responsible for ME in this case.

We can distinguish between two classes of complementarity on the basis of the origin of mutual exclusiveness. These are:

Class I: Pair of variables obtained from a Fourier transform of the state vector and any pair of dynamical variables for which the corresponding operators satisfy a noncommutation relation. These include pairs of canonically conjugate variables.

In this class we have the complementary pairs $\mathbf{r}, \mathbf{p}$ and t, $\mathcal{E}$ derived from $\Psi(r, t)$ and its Fourier transform $\Phi(p, \mathcal{E})$. Also, all pairs of noncommuting variables such as components of angular momentum $J_x, J_y; J_x, J_z$, etc., belong to this class. Similarly, the radiation field variables $\boldsymbol{E}$ and $\boldsymbol{H}$, which satisfy field commutation relation, may serve as the complementary pair in the case of a light beam. Existence of more esoteric complementary pair of variables has also been reported from experiments on elementary particles[14]. ME follows quite clearly from the general quantum-mechanical uncertainty principles for this class of complementary pairs[15].

Class II: A property which depends on interference terms arising from the superposition of a number of states and the property associated with the which-state information form a pair of complementary properties.

In the simplest example of the widely discussed double-slit experiment, one may observe class II type of complementarity by measurement of a single variable. Let the state of a beam of electrons of a definite momentum be given by $\exp(ipz/\hbar)$, the momentum p being along the z direction. At $z = z_1$, a screen is placed with two slits whose centres are at $y = \pm a$ (we take y along the vertical direction). On the right-hand side of the screen the state vector is modified to $\psi = \psi_1 + \psi_2$, where ψ_1 is the wave vector originating from the upper slit and ψ_2 that from the lower slit. If we place a screen in this region at $z = z_2$, we shall observe the interference

Box 3.

Precise statement of BCP and wave–particle duality according to formal quantum mechanics

Existence of complementary pair of variables/properties distinguishes the world of quantum phenomena from the realm of classical physics. Mutual exclusiveness is the *defining property* of complementary pairs.

On the basis of the origin of mutual exclusiveness we can distinguish between two classes of complementary properties. These are:

Class I: Pair of variables obtained from a Fourier transform of the state vector and any pair of dynamical variables for which the corresponding operators satisfy a noncommutation relation. These include pairs of canonically conjugate variables. For this class of complementary pairs, ME follows quite clearly from the general quantum-mechanical indeterminacy principles.

Class II: A property which depends on interference terms arising from the superposition of a number of states and the property associated with the which-state information form a pair of complementary properties. Superposition of states is interpreted with respect to the properties which are measured to show the interference effect and give the which-state information. *This should be considered as the precise expression for the wave–particle complementarity.* The collapse hypothesis of QM ensures mutual exclusiveness in this type of complementarity.

effect. For simplicity we consider ψ_1, ψ_2 as functions of y and z alone. Thus, the probability density function for y as seen on the screen at z_2 is

$$\rho(y, z_2) = |\psi_1(y, z_2)|^2 + |\psi_2(y, z_2)|^2 + \psi_1^*(y, z_2)\,\psi_2(y, z_2) + \psi_2^*(y, z_2)\,\psi_1(y, z_2)$$

The last two terms here give the interference effect. If at $z = z_1$ we make a measurement of y, the result will give us the which-state information and the state vector will collapse to either ψ_1 or ψ_2 and the interference effect in $\rho(y, z_2)$ will disappear. Another interesting feature here is that ψ_1 or ψ_2 need not be eigenfunctions of y; only the respective dispersion in y has to be small.

Experiments based on matter–wave interferometry[2,11,16] belong to this type. A straight edge or a single-slit diffraction experiment may also be included as examples manifesting this type of complementarity. But they seem to have never been discussed in connection with complementarity. The difficulty lies in the fact that the initial state vectors in this case are superposition of a continuous infinity of position eigenfunctions. For example, in the straight-edge case, if we put a series of detectors in the open part near the edge and consider those spots on the screen produced only when the detectors click, the totality of all these spots gives no interference effect. To obliterate the entire pattern, we have to put detectors over a large region near the edge.

Both the interference terms and the which-state information are also obtained from measurement of some suitable pair of variables. Suppose the relevant observables for demonstrating the interference effect is $\hat{\alpha}$ and $\hat{\beta}$. To produce interference effect, the state vector in question must be a superposition of at least two eigenvectors belonging to two distinct eigenvalues of $\hat{\alpha}$ (say). Let us, therefore, take the state vector as $\psi = a_1 |\alpha_1\rangle + a_2 |\alpha_2\rangle$. If we measure the expectation value of $\hat{\beta}$ in this state, we get

$$\langle\hat{\beta}\rangle = |a_1|^2 \langle\alpha_1 |\hat{\beta}| \alpha_1\rangle + |a_2|^2 \langle\alpha_2 |\hat{\beta}| \alpha_2\rangle + a_1^* a_2 \langle\alpha_1 |\hat{\beta}| \alpha_2\rangle + a_2^* a_1 \langle\alpha_2 |\hat{\beta}| \alpha_1\rangle.$$

The last two terms give the interference effect. If $\hat{\beta}$ commutes with $\hat{\alpha}$ then the nondiagonal matrix elements vanish and there will be no interference effect. For this type of experiments it is necessary that $\hat{\beta}$ and $\hat{\alpha}$ correspond to noncommuting variables. Instead of the expectation value, a measurement of the probability distribution of the results of measurement of $\hat{\beta}$ will also reveal the interference effect. If before measuring $\hat{\beta}$ we make a measurement of $\hat{\alpha}$ and the outcome is α_1 (or α_2), the state vector collapses to $|\alpha_1\rangle$ (or $|\alpha_2\rangle$). Subsequent measurements of $\langle\hat{\beta}\rangle$ will give only the first two terms and the interference effect will disappear, showing the complementarity between the interference term in $\langle\hat{\beta}\rangle$ and the which-state information about the state vector. The ME arises here because of the collapse of the state vector and not because of noncommutation between $\hat{\alpha}$ and $\hat{\beta}$.

The simplest example of complementarity of this type is realized in spin interference experiments with neutrons. Let the spin wave function of the neutron beam be given by

$$X = \chi_+ \cos(\theta/2)\, e^{-i\phi} + \chi_- \sin(\theta/2),$$

where $X_\pm$ are the spin eigenstates of $\hat{\sigma}_z$ and the spin of the beam in the state X is polarized along θ, ϕ direction. If we measure the expectation value of $\hat{\sigma}_x$, we get

$$\langle\hat{\sigma}_x\rangle = \sin\theta\, \cos\phi.$$

This is indeed an interference effect. If we measure $\hat{\sigma}_z$ to extract the which-state information, then $\langle\hat{\sigma}_x\rangle$ becomes zero, i.e. the interference effect disappears.

Other examples of class II type of complementarity can be cited from different areas of physics. In class I type of complementarity, the state vector is unimportant. It is the specific pair of variables which satisfy ME irrespective of the state of the system. But in class II type it is the structure of the state vector which is important. Depending upon a specific state vector, one can find a variable (or a pair of variables) and a measurement which will show interference and another measurement which will cause disappearance of interference. The origin of ME for class II type of complementarity is always the quantum-mechanical principle of collapse of a state vector.

So far we have discussed only the extreme situations. Either there is no which-state knowledge and there is complete interference effect or just the opposite. The question arises whether we can quantify the concepts of a degree of interference and a degree of distinguishability of the states. Indeed, using projection operators it is possible to define two operators $\hat{I}$ and $\hat{D}$ which have these properties and whose expectation values satisfy a complementarity relation.

Origin of ME in class II type of complementarity

In a recent paper Uffink and Hilgevoord[17] have made a quantitative analysis of class II type of complementarity. They introduced a measure for the indistinguishability U $(0 \le U \le 1)$ of two quantum states in a given measurement and the amount of interference I observable in the same measurement and established an inequality $U \ge I$, which they regarded as a 'quantitative expression of Bohr's claim that one cannot distinguish

between two possible paths of a particle while maintaining an interference phenomenon'.

In a similar way, but in still simpler terms, we can define I and a distinguishability parameter D (see also ref. 10) with the following considerations: Consider two orthogonal quantum states ψ_1 and ψ_2 expressed in the basis $|i\rangle$ (set of eigenvectors of some quantum-mechanical observable $\hat{\Omega}$) where some eigenvectors are common to both ψ_1 and ψ_2 and others are all different:

$$\psi_1 = \sum_1^n \alpha_i |i\rangle + \sum_{j\neq i} \alpha_j |j\rangle$$

and

$$\psi_2 = \sum_1^n \beta_i |i\rangle + \sum_{l\neq i} \beta_l |l\rangle,$$

where $|j\rangle \neq |l\rangle$ for all j's and l's. In a given measurement, we can define two operators $\hat{D}$ and $\hat{I}$ corresponding to the degree of distinguishability and the amount of interference, observable in the same measurement on a superposed state function ψ as

$$\hat{D} = |\psi\rangle\langle\psi| - \sum_1^n |i\rangle\langle i|, \quad \hat{I} = \sum_1^n |i\rangle|i\rangle.$$

The distinguishability parameter and the amount of interference observable in the same measurement can be obtained from the expectation values of $\hat{D}$ and $\hat{I}$ on the normalized superposed state:

$$\psi = [\psi_1 + \psi_2],$$

where

$$\Sigma|\alpha_i|^2 + \Sigma|\alpha_j|^2 + \Sigma|\beta_i|^2 + \Sigma|\beta_l|^2 = 1;$$

hence,

$$D = \langle\hat{D}\rangle = 1 - \sum_1^n (|\alpha_i|^2) + |\beta_i|^2).$$

So, D lies between 0 and 1. The zero value indicates complete lack of distinguishability and 1 complete distinguishability. A value of 0.6 for D signifies 60% distinguishability, implying that if $\hat{\Omega}$ is measured N times in 60% cases, we shall be able to distinguish the state and in 40% cases we fail,

Similarly,

$$I = \langle\hat{I}\rangle = \sum_1^n (|\alpha_i|^2 + |\beta_i|^2)$$

and, hence, $D + I = 1$.

When $D = 1$, $I = 0$, and vice versa; thus, D and I are complementary aspects. As an illustration consider the case of scattering of two identical particles (i.e. indistinguishable and $D = 0$); here interference of amplitude in the differential scattering cross-section is observed ($I = 1$). But for two distinguishable particles, no interference effect is observed. However, in the case of a double-slit experiment $D = 1$ does not necessarily mean that we have the which-slit information unless we make a measurement. $D = 1$ means that we have, in principle, a 100% efficient which-slit detector for the particular experiment.

We also note that the commutation relation satisfied by the two operators D and I is

$$[\hat{D}, \hat{I}] = [|\psi\rangle\langle\psi| \sum_1^n |i\rangle\langle i| - \sum_1^n |i\rangle\langle i|\psi\rangle\langle\psi|]$$
$$= \sum_1^n [(\alpha_i^* + \beta_i^*)|\psi\rangle\langle i| - (\alpha_i + \beta_i)|i\rangle\langle\psi|]$$
$$\Rightarrow \langle\psi|[\hat{D}, \hat{I}]|\psi\rangle = 0.$$

So the commutation relation does not lead to any uncertainty type of relation between ΔD and ΔI, representing the corresponding dispersions.

In all so-called wave–particle complementary experiments involving one particle at a time, time is an important parameter. We can introduce time in the above analysis and discuss the possibility whether we can have at $t = 0$, $D = 1$ and $I = 0$ and at a subsequent instant in the interference zone $D = 0$ and $I \neq 0$. A double-slit experiment closely resembles this situation. The essence of wave–particle complementarity problem is whether we can make a which-state measurement at $t = 0$ (or in the noninterfering region in the vicinity of the slits) and still get interference at a subsequent instant of time t? This is ruled out in quantum mechanics because of the collapse hypothesis.

Here we note that the operators $\hat{D}$ and $\hat{I}$ are state-dependent and the quantification of the qualitative properties of distinguishability and interference phenomenon is brought by $D + I = 1$. In complementarity phenomena the extremes, i.e. $D = 1$, $I = 0$ or $D = 0$, $I = 1$ are considered. But here the whole range is included. From the foregoing discussions it is clear that class II complementarity is consistent with the principles of QM. ME follows from collapse hypothesis of QM and in that sense we may say that class II BCP is derived from QM.

With this formulation we conclude that any experiment which claims to violate ME in wave–particle complementarity will imply some limitations of the quantum-mechanical formalism itself. This we believe is in true spirit of Bohr's original ideas.

In the following section we discuss some significant experiments on wave–particle duality in the light of the present formulation. All these experiments may be broadly

classified into three classes. The first group of experiments[2,11,18,19] are claimed to provide experimental evidence in favour of ME in wave–particle duality. In a rather novel thought experiment suggested by Scully *et al.*[11], the incompatibility between 'welcher weg' (which path) knowledge and the appearance of interference cannot be traced to any uncertainty relation. On the other hand, certain 'welcher weg' experiments[20–22] which made joint unsharp measurements of complementary observables have been used to argue that they yield 'partial particle' and 'partial wave' information and are, therefore, in conflict with the Bohrian notion of ME.

Recently, Petroni and Vigier[16] and Ghose *et al.*[23] have proposed experiments to demonstrate that a quantum-mechanical entity may display its particle and wave aspects at the same time, thereby contradicting the principle of complementarity. In what follows the significance and import of these experiments are analysed in the light of the precise statement of BCP.

Discussion of some experiments

We now take a brief look at the well-known thought experiments on wave–particle duality. In the archetypal example of the recoiling-slit arrangement Einstein hoped to devise a gedanken experiment to provide the which-path information without destroying the interference pattern of the photons; Bohr[18] pointed out that we must also treat the recoiling slits by the laws of QM. Following the arguments of Feynman[2], we note that in order to observe the interference pattern, the slits must be localized with an accuracy of the order of the slit width and the corresponding uncertainty in the slit momentum is always much greater than the momentum transfer in the scattering process. Consequently, the which-path information is not available when the photons do produce interference pattern. The entire analysis here envisages a classical particle-like passage of the photons (even when they do interfere!) through the double slit, while a correct quantum-mechanical explanation should be obtained in terms of the quantum-mechanical field function satisfying appropriate boundary condition at the slits – the concept of momentum transfer being out of place[15]. However, in the similar example of double-slit interference experiment with electrons, when Feynman intends to 'watch' the electron using scattered photon to do the spying, the interference pattern gets washed out, and Feynman's analysis invoking the uncertainty principle for the scattering process in this case may be regarded as an intuitive description not quite consistent with the formalism of QM[15].

In the interference region, the wave function describing the interfering beams may be written as the sum of two terms referring to the two slits:

$$\Psi(r) = [\Psi_1(r) + \Psi_2(r)]$$

The corresponding probability density at $r = R$ on the screen, denoted by $P(R)$, will be given by the squared modulus of $\Psi(R)$ and contains the interference term represented by $\Psi_1^* \Psi_2 + \Psi_2^* \Psi_1$. Now, if we consider two detectors 1 and 2 in the vicinity of the two slits, either making transition from ground state g to an excited state e induced by the passage of the interfering quantum object, the state of the beam plus the detector system (after entanglement) may be given by

$$\Psi(r) = [\Psi_1(r)\ |1_e 2_g\rangle + \Psi_2(r)\ |1_g 2_e\rangle].$$

For nonorthogonal detector states, partial or complete interference phenomenon is observed and no which-state information is available. In some special cases, as for detectors having coherent-state-like energy spectrum, one can have the so-called 'ein weg' (one way) knowledge, as claimed by Petroni and Vigier[16]. But if the detector states are orthogonal, the entanglement with the detector states results in a loss of coherence in interferometer and we get no interference. However, from an appropriate choice of the detector basis, as in the case of Scully *et al.*'s[11] micromaser detector, possibility of quantum erasers appears naturally[24] and the interference pattern can be retrieved. On the other hand, from a subsequent measurement on the detector state (with the aid of an atomic reader as in Scully's example) the superposed state gets collapsed and we get the which-state information.

The experiment of Petroni and Vigier

In the proposed neutron interferometry experiment[16,25] a pulsed neutron beam polarized vertically with respect to the apparatus is incident on a perfect crystal interferometer. At any instant, there is at most one neutron inside the interferometer. The beams from the beam splitter together with the coil states, described by the wave function

$$\Psi(r) = [\Psi_1(r) + \Psi_2(r)]\ |_\uparrow\rangle \ o\ |C_1 C_2\rangle,$$

are then made to pass through two radio-frequency (rf) spin flip coils, one in each arm, and are superposed beyond the final set of crystal planes.

The wave function of the emerging beams from the coils is given by[26]

$$\Psi(r) = [\Psi_1(r)\ |C_1' C_2\rangle + \Psi_2(r)\ |C_1 C_2'\rangle\ ]\ |^\downarrow\rangle$$

and retains the coherent superposition as the coil states (denoted by the C's) are nonorthogonal. The final beam thus produces an interference in the measured intensity.

There are no resonant frequencies provided by the rf field which correspond to half the energy of spin flip. If we assume that energy and momentum are conserved in each energy exchange, it implies that a spin flip occurs in either of the coils and not in both. It is also possible to obtain interference in the intensity of the detected neutrons in the same experimental set-up. However, as pointed out by Scully and Walther[27] and also by Unnerstall[26], the rf coils generate coherent states of photons and the passage of neutron through the coil cannot leave the which-path information in the coil since the coherent photon distribution remains essentially unchanged by the addition of a single photon associated with spin flip. After the passage of a large number of neutrons, the energy transferred to each coil is summed up to an amount that is detectable, and one can tell that out of N neutrons involved in the experiment, n (say) neutrons have exchanged energy with coil I and the rest $N-n$ with coil II. This is precisely what Petroni and Vigier[16] claimed as 'ein weg' information. It seems legitimate to realist physicists to infer that each neutron followed a particular trajectory or the other through the rf coils, leaving the so-called 'ein weg' information in the process. But this extrapolation is, however, not endorsed by quantum orthodoxy[26]. 'Ein weg' experiment does not contradict BCP. The only problem it creates is how one would accommodate the 'ein weg' information within the framework of the orthodox interpretation of QM.

Experiment of Ghose et al.[23]

Ghose *et al.*[23] have proposed an experiment in which single-photon states are incident on a combination of two prisms placed opposite to each other. Mizobuchi and Ohtake[28] have performed the actual experiment and verified the quantum-optical predictions outlined in the former paper. When the gaps between the prisms is larger than the wavelength, the incident 'photon states' suffer total internal reflection inside the first prism (registered by counter 1). When the gap is shorter than the wavelength, there is a possibility of their 'tunnelling' across the gap (registered by counter 2).

Since they consider the presence of 'tunnelling' to be accounted for by assigning a wave aspect to the photon, the perfect anticoincidence in the clicking of the two counters led them to claim that the 'experimental result contradicts the tenet of mutual exclusiveness of the classical wave and particle pictures assumed in Bohr's complementarity principle'.

We[23] have adopted the viewpoint of Grangier *et al.*[19] and consider anticoincidence as a signature of particle property. However, the anticoincidence observed in two detectors in a single-photon field has a totally different explanation in quantum field theory[25], devoid of the photon particle picture, a photon being just an excitation of a mode.

On the other hand, the so-called 'tunnelling' cannot be accepted as a hallmark of wave property. In that case we should consider the tunnelling of α-particle through atomic nucleus as a violation of ME in wave–particle dualism. Furthermore, while Ghose *et al.* consider the existence of evanescent waves leading to transmission of light at the critical angle as a specific wave property, Chiao *et al.*[29] described the phenomenon of 'frustrated total internal reflection' of the purely nonclassical single-photon state exploiting an analogy with particle (quantum) tunnelling through the classically forbidden region. They have also proposed an experiment to measure the photon tunnelling time using the 'particle aspect' of photon tunnelling.

According to Chiao *et al.*[29], there exists a classical limit associated with coherent states of the electromagnetic field in which the tunnelling phenomenon can be understood entirely as a classical wave phenomenon. The two arguments are, therefore, irreducible. However, it seems that tunnelling can be interpreted both as a particle and a wave property.

If we accept the interpretations of Grangier *et al.*[19] and Ghose *et al.*[23] of wave–particle duality in BCP then Mizobuchi and Ohtake's experiment[28] definitely contradicts ME. Similarly, a measurement of photon velocity on refraction in a denser medium coupled with anticoincidence will also violate ME. Again, since the probability of joint detection of more than one photon is exactly zero for an 'ideal single-photon' state, should we claim that both 'particle' and 'wave' appear in the single run of an interference experiment with 'ideal single-photon' states? One can multiply this sort of examples indefinitely[30]. On this point it is also important to note that photon enters the theory as a secondary quantity – an elementary excitation of the field. The concept of photon does not presuppose that the quantum of field energy is concentrated at a point in space, i.e. that the photon is a localized entity analogous to a very small bullet. The photon is only approximately localizable and the wave function in coordinate space is nonlocal[1]. We may also recall that Einstein always stressed the provisional nature of the light quantum postulate. In December 1951, he wrote to his friend Besso: 'All these fifty years of pondering have not brought me any closer to answering the question "What are light quanta?". Nowadays, every Tom, Dick and Harry thinks he knows it, but he is mistaken'[31]. Pushing the classical analogy too far in the photon picture may, therefore, lead to contradictions.

Some physicists are, in effect, asserting that any property for which a standard classical corpuscular theory does not exist is a wave property and vice versa. This rather simplistic interpretation of Bohr's idea is un-

tenable, for a 'particle' is not always a classical particle (when it tunnels!) and similarly a 'wave/field' may have nonclassical features. *In fact, the use of the term wave–particle duality is at the root of much confusion and QM does not guarantee the claim of complementarity between arbitrarily defined wave and particle properties.*

In this experiment the pair of properties involves no superposed state and hence no interference effect is there. Therefore, it belongs neither to class I nor to class II types of BCP. We conclude that this experiment is not suited to test BCP in its precise version.

The unsharp joint measurements[20–22] which claim to observe both interference and the path of the particle do not really fall within the compass of BCP. Bohr himself in the Como lecture discussed the wave packet description of a particle (where approximate knowledge of both momentum and position exists) as an illustration of the complementarity principle, implying that unsharp knowledge of conjugate variables does not violate BCP. Both class I and class II types of BCP imply sharp measurement.

Finally, let us summarize the situation. If we interpret the wave–particle duality aspect of BCP as a principle independent of the formalism of QM, then mutual exclusiveness becomes a doubtful proposition. In the quantum region we have neither a true classical particle nor a true classical wave. In some experiments, for some limiting value of a parameter, results can be interpreted as those of a classical particle and in the other limit as those of a classical wave. Naturally, for any intermediate value of the parameter, this type of exclusive interpretation would be impossible. Moreover, to try to designate an entity as either a particle or a wave by measuring a single property is not a very satisfactory and unambiguous procedure. A classical particle or a wave is defined by a totality of many properties.

On the other hand, we can look at BCP as focusing our attention on an important and characteristic feature of QM that every physical entity has a number of pairs of mutually exclusive variables/properties which is unknown in classical physics. Then ME becomes a defining feature of BCP. Selection of pairs of such variables/properties has to be made consistently with QM. This is really possible, as has been shown in this paper. With this modified formulation, BCP becomes consistent with QM and violation of BCP without violating QM becomes impossible.

1. Akhiezer, A. I. and Berestetskii, V. B., *Quantum Electrodynamics*, Interscience, New York, 1965, pp. 4–5.
2. Feynman, R. P., Leighton, R. B. and Sands, M., *The Feynman Lectures on Physics*, Addison-Wesley/Narosa, New Delhi, 1988, vol. III, pp. 1–6.
3. Bohm, D., *Wholeness and the Implicate Order*, Routledge & Kegan Paul, London, 1980, p. 128.
4. Gudder, S. P., in *The Wave–Particle Dualism – A Tribute to L. de Broglie on his 90th Birthday* (eds Diner, S., Frague, D., Lochak, G. and Selleri, F.), Dordrecht, 1984, p. 499.
5. Popper, K. R., in *Quantum Theory and Schism in Physics* (ed. Bartley, W. W.), Hutchinson, London, 1982, p. 126.
6. Moore, W. J., *Schrödinger – Life and Thought*, Cambridge University Press, Cambridge, 1989, p. 187.
7. Bohr, N., Como Lecture 1927, reprinted in *Nature*, 1928, **121**, 580–590.
8. Schilpp, P. A. (ed.), *Einstein: A Philosopher–Scientist*, Library of Living Philosophers, Evanston, IL, Harper and Row, New York, 1949, vol. 7, pp. 663–688.
9. Bohr, N., *Atomic Theory and Description of Nature*, Cambridge University Press, 1934, p. 10.
10. Sen, D., Basu, A. N. and Sengupta, S., *Helvetica Physica Acta*, 1994, **67**, 785–798.
11. Scully, M. O., Englert, B. G. and Walther, H., *Nature*, 1991, **351**, 111–116.
12. Bacry, H., *Localizability and Space in Quantum Physics*, Springer, Berlin, 1988, pp. 20–21.
13. Bohr, N., *The Causality Problem in Atomic Physics*, International Institute of Intellectual Cooperation, Paris, 1939, p. 24.
14. Krisch, A. D. *et al.*, University of Michigan Preprint UM HE 85-17 (1985) and reported in an article by Kursunoglu, B. N. in *Reminiscences About a Great Physicist, Dirac, P. A.* (eds Kursunoglu, B. N. and Wigner, E. P.), Cambridge University Press, Cambridge, 1987, p. 289.
15. Sen, D., Basu, A. N. and Sengupta, S., *Z. Naturforsch. A.*, 1994, communicated.
16. Petroni, N. C. and Vigier, J. P., *Found. Phys.*, 1990, **22**, 1–40; Rauch, H. and Vigier, J. P., *Phys. Lett.*, 1990, **A151**, 269–275.
17. Uffink, J. and Hilgevoord, J., *Physica*, 1988, **B151**, 309–313.
18. Bohr, N., in ref. 8, pp. 200–241; reprinted in *Quantum Theory and Measurement* (eds Wheeler, J. A. and Zurek, W. H.), Princeton University Press, Princeton, 1983, pp. 9–49.
19. Grangier, P., Roger, G. and Aspect, A., *Europhys. Lett.*, 1986, **1**, 173–179; Aspect, A. and Grangier, P., *Hyperfine Interact.*, 1987, **37**, 3–17.
20. Wooters, W. K. and Zurek, W. H., *Phys. Rev.*, 1979, **D19**, 473–484.
21. Mittelstaedt, P., Prieur, A. and Schieder, R., Proc. Symp. Found. Mod. Phys., Finland (eds Lahti, P. and Mittelstaedt, P.), World Scientific, Singapore, 1987, p. 403–418.
22. Bartell, L. S., *Phys. Rev.*, 1980, **D21**, 1698–1699.
23. Ghose, P., Homd, D. and Agarwal, G. S., *Phys. Lett.*, 1991, **A153**, 403–406; Ghose, P., Home, D. and Agarwal, G. S., *Phys. Lett.*, 1992, **A168**, 95–99.
24. Tan, S. M. and Walls, D. F., *Phys. Rev.*, 1993, **A47**, 4663–4676.
25. Dewdney, C., Horton, G., Lam, M. M., Malik, Z. and Schmidt, M., *Found. Phys.*, 1992, **22**, 1217–1265.
26. Unnerstall, T., *Phys. Lett.*, 1990, **A151**, 263–268.
27. Scully, M. O. and Walther, H., *Phys. Rev.*, 1989, **A39**, 5229–5236.
28. Mizobuchi, Y. and Ohtake, Y., *Phys. Lett.*, 1992, **A168**, 1–5.
29. Chiao, R. Y., Kwiat, P. G. and Steinberg, A. M., *Physica*, 1991, **B175**, 257–262.
30. Rangwala, S. and Roy, S. M., *Phys. Lett.*, 1994, **A190**, 1–4.
31. Einstein, A. and Besso, M. in *Correspondences 1903–1955* (ed. Speziali, P.). Herman, Paris, 1972, p. 453; also quoted by Pais, A., in *Subtle is the Lord, The Life and Science of Einstein, A.*, Oxford University Press, 1982, p. 382.

ACKNOWLEDGEMENT. We thank the anonymous reviewer for positive comments and constructive criticism.

Received 6 February 1995; revised accepted 18 July 1995

RESEARCH ARTICLE

α-Turns in protein structures

D. V. Nataraj, N. Srinivasan*, R. Sowdhamini* and C. Ramakrishnan†

Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012, India
*Present address: ICRF Unit of Structural Molecular Biology, Department of Crystallography, Birkbeck College, London, WC1E 7HX, UK

The occurrence of $5 \rightarrow 1$ type of hydrogen bonds (α-turn) in proteins has been studied using a data set comprising of 107 proteins with resolution ≤ 2.0 Å. A very large majority of such α-turn segments (96%) form part of regular α-helices. The examples (84) which do not form part of an α-helix are termed 'isolated α-turns' and are grouped into two major families and seven minor groups along with two isolated examples based on the similarity of conformational angles. The family with large number of examples (50) have (φ, ψ) angles close to an α-helix and hence belong to the class of the shortest α-helices. The 'end to end' distances of these α-turns vary between 4.7 and 6.7 Å, the range being nearly the same as that of α-helices. The propensity calculations show that some amino acids such as Glu, Ser and Thr have statistically significant higher preferences to occur in α-turns than in α-helices. In addition to the $5 \rightarrow 1$ type, the residues in the α-turn are involved in hydrogen bonds with other parts of the chain. The residues are in general more hydrophilic compared to those in α-helices. In many cases (70%) the α-turn occurs at the ends of extended strands, and whenever it occurs at the loop regions connecting two extended strands, it brings about a hairpin bend.

POLYPEPTIDE chain reversal is an essential feature in bringing about the globular shape of proteins. Turns are also associated with various aspects of protein function[1-3]. The direction of the main chain can be reversed by characteristic conformations at a single amino acid residue (γ-turns[4,5]) or at two residues (β-turns[2,6,7]) or by longer loops[8-10]. Both γ- and β-turns are locally stabilized by hydrogen bonds in the main chain. The $3 \rightarrow 1$ hydrogen bond in the γ-turn is formed between carboxyl oxygen at position i and amide at position $i+2$, with the turn forming residue at the position $i+1$. Most β-turn types are stabilized by a $4 \rightarrow 1$ hydrogen bond between carbonyl oxygen at position i and amide at position $i+3$.

These two kinds of turns occur frequently in globular proteins, with β-turn being one of the more common substructures[11-14]. In this communication, attention is drawn to 'three-residue segments' locally stabilized by a $5 \rightarrow 1$ hydrogen bond between carbonyl oxygen at position i and amide at position $i+4$, reversing the course of the polypeptide direction. The analysis of a large number of known three-dimensional structures of proteins reveals that such a local chain reversal is a common occurrence. This class of turns is referred to as α-turns since the hydrogen bonding pattern is the same as in the α-helix. A schematic diagram of an α-turn is shown in Figure 1. The present analysis shows that α-turns can have different characteristic main chain conformational angles and amino acid preferences. They connect regular secondary structures and quite often act as linkers of two β-strands.

†For correspondence.

Materials and methods

Data set used

The data set used in the present study consists of a set of 107 proteins whose coordinates are available in Protein Data Bank[15,16]. The choice of the set has been guided by the condition that their resolution is 2 Å or better, and the cut-off percentage of the sequence homology between any two proteins in the set is not more than 40. Also in the case of proteins having multiple chains with the same sequence, only one chain has been considered.

Identification of α-turns in proteins

Since the $5 \rightarrow 1$ hydrogen bond is the crux of the

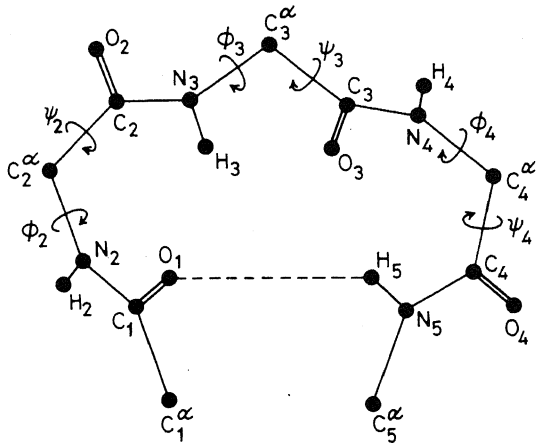


Figure 1. Schematic diagram of an α-turn, occurring in a system of four linked peptide units. The designations of atoms and torsion angles involved are also marked.

α-turn, the proteins are examined to find out all tetrapeptide segments which have this hydrogen bond. The hydrogen bond criteria used for this purpose is as follows.

The hydrogen bond length l (N . . O distance) must lie in the range 2.6–3.6 Å and the hydrogen bond angle θ (H–$\hat{N}$ . . O) should be less than or equal to 40°. (The criteria used here are slightly relaxed compared to those that have been normally used in other conformational studies. For example, in small molecules, where the crystal structures are accurately determined, the hydrogen bond length is in the range 2.6–3.2 Å and hydrogen bond angle is usually less than 30° (ref. 17)). The reason for relaxing the criteria is the lesser accuracy in protein structure determination as well as the lack of information on the positional coordinates of hydrogen atoms.)

As the hydrogen positions are not available in the coordinate data of Protein Data Bank (PDB) (with few exceptions, however), the H-bond angle θ could not be readily evaluated. For this purpose, the positional coordinates of the amide hydrogens have been computed by geometrically fixing the hydrogen atoms at the corresponding nitrogen atoms in the plane of the peptide unit using the ideal bond angles around the nitrogen atom[18].

Results and discussion

Using the H-bond criteria mentioned above, it has been possible to identify 2525 tetrapeptide segments in the proteins which have $5 \rightarrow 1$ hydrogen bonds. These examples obviously include those segments which are part of α-helices, since an α-helix is a succession of $5 \rightarrow 1$ hydrogen-bonded segments. The removal of the segments that are part of an α-helix has been done by a combination of (φ, ψ) criteria and by inspection. The conformations at (φ, ψ) at the five α-carbon atoms are designated as R or X, whether or not it lies in the range -130 to 0° for φ and -130 to 30° for ψ. If there are four **successive** R's (i.e. the designations at the four α-carbon atoms i, $i+1$, $i+2$ and $i+3$ or at $i+1$, $i+2$, $i+3$ and $i+4$ are R's) then it is treated as forming a part of α-helix. The examples which still remain are further examined by inspection to remove any possible borderline cases. On removing those examples which form part of an α-helix, there were 84 examples which remained, and these can be considered as isolated α-turns. (The corresponding numbers of examples for β- and γ-turns are 383 and 210, respectively). The PDB codes of the proteins and the segments in which these α-turns occur are listed in Table 1 *a*. These examples occur in 51 of 107 proteins of the total set. The number of examples of these α-turns in a protein varies from one to four except in the protein phosphate dehydrogenase (1gd1), where there are five examples.

Table 1a. α-Turn segments and the residues in them occurring in the data set used, grouped into different families and groups.The number of examples in each family/group is given within parentheses

PDB code	Segment	Residues				
Family F1 (50)						
1aap	A24–28	D	V	T	E	G
1acx	82–86	D	C	A	T	D
1ak3	A149–153	D	D	L	T	G
1cox	391–395	N	A	G	S	G
1drf	152–156	D	L	E	K	Y
1eca	38–42	F	T	Q	F	A
1gd1	O47–51	D	S	V	H	G
1gd1	O138–141	D	P	K	A	H
1gox	345–349	S	L	K	E	I
1mba	43–47	F	A	D	F	K
1omd	2–6	I	T	D	I	L
1ova	A277–281	S	S	N	V	M
1ova	A318–322	L	S	G	I	S
1rdg	6–10	C	T	V	C	G
1rdg	14–18	D	P	A	K	G
1rdg	39–43	C	P	V	C	G
1sn3	7–11	K	K	S	D	G
1sn3	31–35	A	K	N	Q	G
1thb	A113–117	L	P	A	E	F
2act	85–89	T	E	E	N	Y
2alp	140–156	G	R	T	T	G
2aza	A40–44	A	K	S	A	M
2ca2	34–38	D	T	H	T	A
2cpp	77–81	D	Y	R	H	F
2cpp	328–332	N	E	R	E	N
2csc	59–63	D	P	D	E	G
2cyp	58–62	D	K	H	D	N
2er7	E240–244	S	S	S	V	G
2fcr	94–98	D	A	E	G	Y
2fcr	148–152	D	M	V	N	D
2ltn	A54–58	D	R	E	T	G
2ltn	A125–129	N	A	A	W	D
2ltn	A167–171	N	A	A	T	N
2rhe	93–97	N	D	S	L	D
2rsp	A46–50	S	E	E	D	W
2sga	140–156	G	S	T	T	G
2sga	219–223	N	C	R	T	G
2trx	A59–63	N	I	D	Q	N
3blm	50–54	D	T	K	S	G
3cla	97–101	H	Q	E	T	E
3cla	194–198	H	H	A	V	C
3fgf	67–71	G	V	C	A	N
3grs	164–168	H	E	S	Q	I
4cha	A95–99	N	S	L	T	I
4enl	102–106	K	S	K	L	G
4fd1	35–39	H	P	D	E	C
5cpa	3–7	S	T	N	T	F
5cpa	29–33	H	P	E	L	V
5p21	145–149	S	A	K	T	R
6ldh	181–185	H	S	C	S	C
Family F2 (9)						
1ak3	A137–141	N	I	E	F	N
1cox	453–457	L	L	N	S	A
1fkf	87–91	H	P	G	I	I
1gdl	O129–133	V	M	G	V	N
1gdl	O267–271	L	K	G	I	L
2tec	E261–265	G	T	G	T	Y
2trx	A49–53	Y	Q	G	K	L
4fdl	8–12	C	I	K	C	K
5cpa	89–93	N	Y	G	Q	N
Group g1 (5)						
1gcr	47–51	R	P	N	Y	Q

Table 1a (*contd.*)

PDB code	Segment	Residues				
1gcr	136–140	M	P	S	Y	R
2cyp	35–39	E	Y	D	N	Y
2fbj	L48–52	Y	E	I	S	K
2prk	212–216	W	I	G	G	S
Group g2 (2)						
2fbj	H100–104	H	Y	Y	G	Y
2rhe	50–54	Y	Y	N	D	L
Group g3 (4)						
1fkf	82–86	Y	G	A	T	G
1ova	A69–73	K	D	S	T	R
3grs	55–59	G	G	T	C	V
4bp2	25–29	Y	G	C	Y	C
Group g4 (4)						
2act	188–192	W	G	E	E	G
2lzm	27–31	I	G	I	G	H
3apr	E11–15	Y	G	N	D	I
4fxn	56–60	M	G	D	E	V
Group g5 (3)						
1fxl	71–75	F	I	P	L	F
1ypi	A25–29	E	R	L	N	T
4fxn	40–44	I	D	E	L	L
Group g6 (3)						
1gdl	O300–304	I	D	G	K	M
2ltn	A100–104	Y	L	G	V	F
4pep	9–13	Y	L	D	T	E
Group g7 (2)						
1ak3	A129–133	H	P	G	S	G
1fxl	72–76	I	P	L	F	D
Isolated examples						
1rbp	63–67	L	L	N	N	W
2fbj	L166–170	D	S	K	D	S

In those proteins which contain two or more examples, the segments are well separated from one another. The two examples occurring in protein flavodoxin (1fx1) are somewhat unique in that the two α-turn segments 71–75 and 72–76 overlap, resulting in a longer segment 71–76. The individual segments are, however, conformationally different and the full segment 71–76 is not α-helical.

For those segments that have been picked up as α-turns, the (φ, ψ) angles at the three intervening α-carbon atoms C_2^α, C_3^α, and C_4^α have been computed and analysed.

Conformational analysis

The (φ, ψ) angles at each of the three C^α atoms have been plotted on the (φ − ψ) plane and these are shown in Figure 2 *a–c* corresponding to C_2^α, C_3^α, and C_4^α, respectively. These plots are first analysed independently and then used collectively to arrive at families of α-turns.

In Figure 2 *a*, which corresponds to the conformation at C_2^α, four groups of points are plainly seen, one of them being a dense cluster having a very large number of examples near the right-handed α-helical region designated as α_R. Of the remaining three clusters, one which occurs in the left-handed α-helical region is designated as α_L. The third, having few examples, occurs near the extended conformational region and has been designated as E. The fourth cluster which has eight examples occurs in a region that can be considered as inverse to the cluster E and has been denoted as E′. (Seven out of these eight examples are Gly and the odd member out is the residue Asp, A70 in the protein ovalbumin[19]

Table 1b. Average (φ, ψ) values (rounded off to the nearest 5°) at the three intervening α-carbon atoms of the α-turn segment along with their designations

Family/ group	Average (φ, ψ) at C_2^α	C_3^α	C_4^α	Designation*
F1	(−65, −25)	(−75, −35)	(−100, −20)	$\alpha_R-\alpha_R-\alpha_R$
F2	(−60, 130)	(75, 0)	(−115, −50)	$E-\alpha_L-\alpha_R$
g1	(−65, 155)	(70, 30)	(60, 40)	$E-\alpha_L-\alpha_L$
g2	(0, 95)	(70, −15)	(−40, 25)	$\alpha_L-X-\alpha_R$
g3	(60, −145)	(−75, −25)	(−80, −40)	$E'-\alpha_R-\alpha_R$
g4	(60, −140)	(−95, 75)	(60, 40)	$E'-P-\alpha_L$
g5	(−50, −60)	(−70, 105)	(120, −50)	α_R-P-X
g6	(50, 30)	(60, 10)	(−120, −30)	$\alpha_L-\alpha_L-\alpha_R$
g7	(−65, 110)	(150, −60)	(−80, −30)	$E-X-\alpha_R$
#	(−50, −40)	(−60, −15)	(80, 45)	$\alpha_R-\alpha_R-\alpha_L$
#	(−105, 45)	(−130, −40)	(−100, −30)	$B-\alpha_R-\alpha_R$

*For the location of regions α_R, α_L, E, E′, B, P, and X see Figure 2.
– Isolated examples.

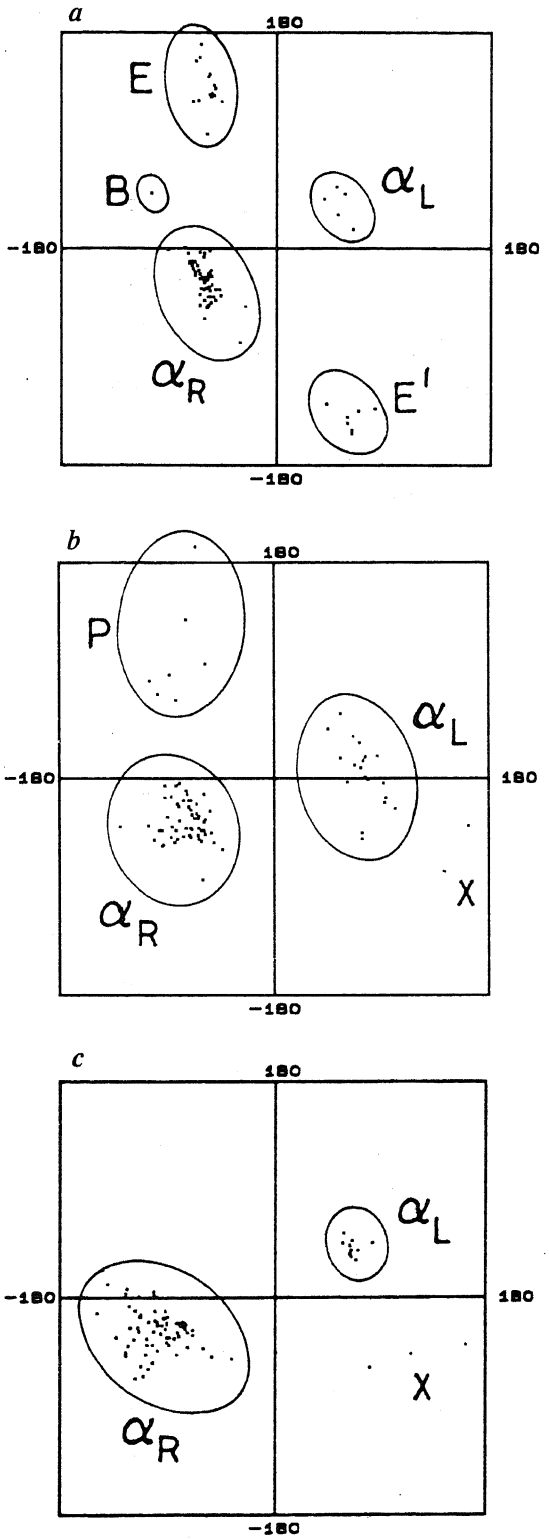


Figure 2. The (ϕ, ψ) plot of the conformational points at the three α-carbon atoms in the examples of isolated α-turns: *a*, at C_2^α; *b*, at C_3^α; *c*, at C_4^α. The clusters in each of the plots are also marked, with their designations indicated.

(1ova)). In addition to this, there is a solitary example at (– 106°, 46°) (referred to as B, as this occurs in the bridge region of ($\phi - \psi$) map[20]). Thus, the distribution at C_2^α points to the possibility of at least four major families.

Figure 2 *b*, which corresponds to the ($\phi - \psi$) distribution at C_3^α, shows three prominent clusters. The one which lies in the α_R region comprises a majority of the points. The other two clusters are slightly more scattered as compared to the α_R region. One of these occurs in the right half of the map spanning across the ϕ-axis. This is denoted as α_L. The third one occurs in the left top quadrant of the map, in which the extended conformations normally occur. But five of the seven points are clustered closer to the poly proline region of the map (– 60°, 120°) than the $\psi = 180°$ axis. Hence, we have preferred to use the symbol P rather than E to denote this cluster (see Figure 2 *a*). In addition to these, there are two isolated points occurring near the right edge of the map which are designated as X.

Figure 2 *c*, which shows the distribution of (ϕ, ψ) at C_4^α, has two clusters, a major one in α_R and a minor one in α_L region. In addition to these there are three isolated points (X) occurring in the right bottom quadrant of the map. It may be worthwhile to note that, there are no points in the 'extended' region of the map as contrasted with Figure 2 *a* and *b*.

Classification into families

As the formation of $5 \rightarrow 1$ hydrogen depends on the *combined set* of (ϕ, ψ) values at the three α-carbon atoms *in a collective way*, the examples have been carefully looked into and grouped into different families such that members of the same family have nearly the same combination of (ϕ_2, ψ_2), (ϕ_3, ψ_3), (ϕ_4, ψ_4). This exercise has revealed that a large fraction of examples (60%) form one major family (designated as F1). The other families contain relatively small number of examples. The family which has the next highest number (9 examples) is designated as F2. Apart from these, seven more groups can be identified, having 2–5 examples. In view of the small number of examples, these will be referred to as groups g1–g7. In addition, there are two isolated examples. The details pertaining to the members in various families and groups are given in Table 1 *a*. The (ϕ, ψ) plots are shown in Figure 3 *a–k* and the average (ϕ, ψ) values at C_2^α, C_3^α and C_4^α for each category are given in Table 1 *b*.

The family F1, which has a bulk of examples, is characterized by α_R conformations at all the three carbon atoms C_2^α, C_3^α and C_4^α. Thus, this family, which can be designated as 'α-helical α-turn' represents the shortest helix that can possibly occur. The (ϕ, ψ) values at the two α-carbon atoms immediately preceding and succeeding the turn, i.e. at C_1^α and C_5^α, are shown in Figures 4 *a* and *b* (absence of points in the α_R region in this figure confirms the isolated nature of the α-turn in these

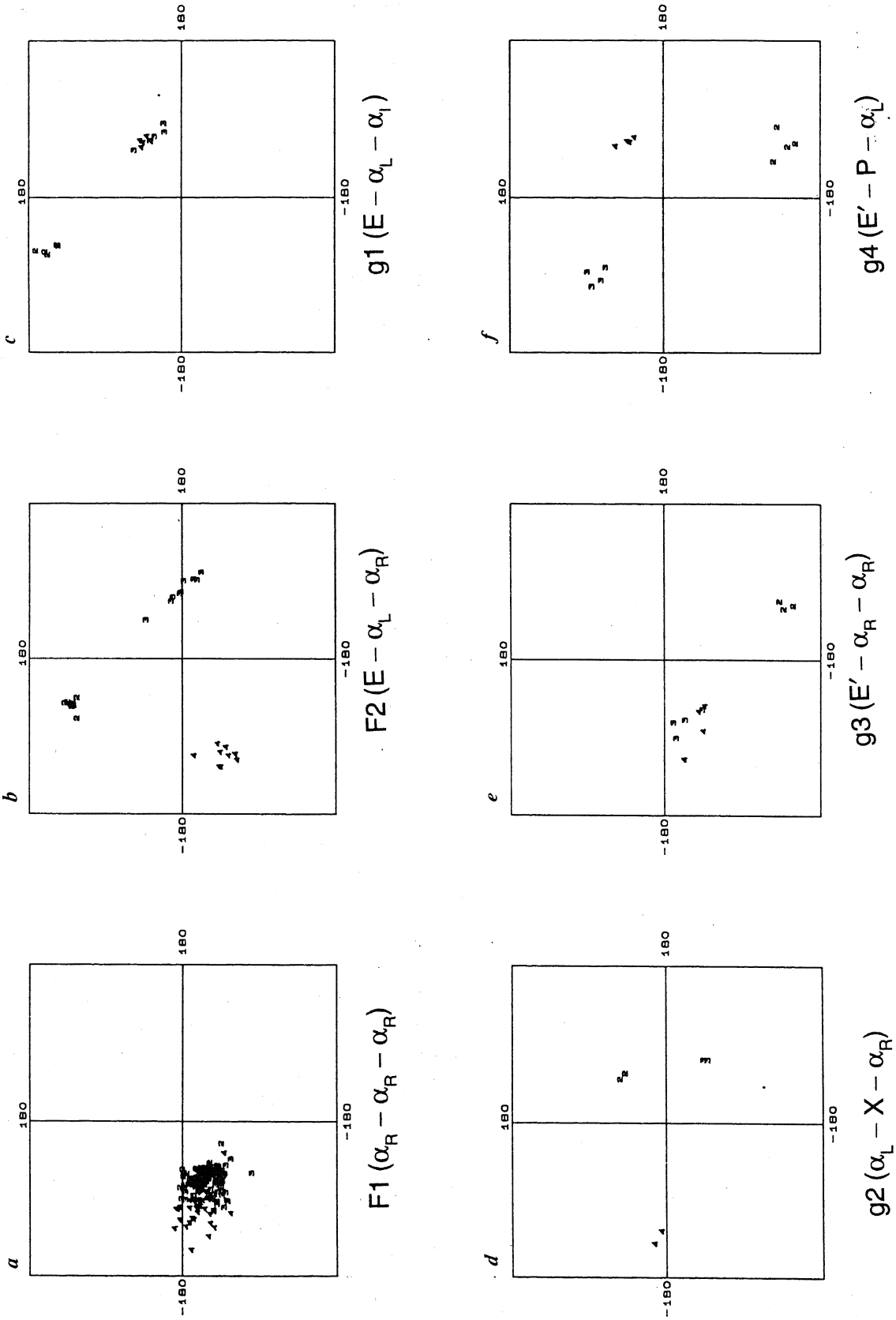

a
b
c
d
e
f
180
-180
F1 (αR − αR − αR)
F2 (E − αL − αR)
g1 (E − αL − αl)
g2 (αL − X − αR)
g3 (E′ − αR − αR)
g4 (E′ − P − αL)

Figure 3 *a–f.*

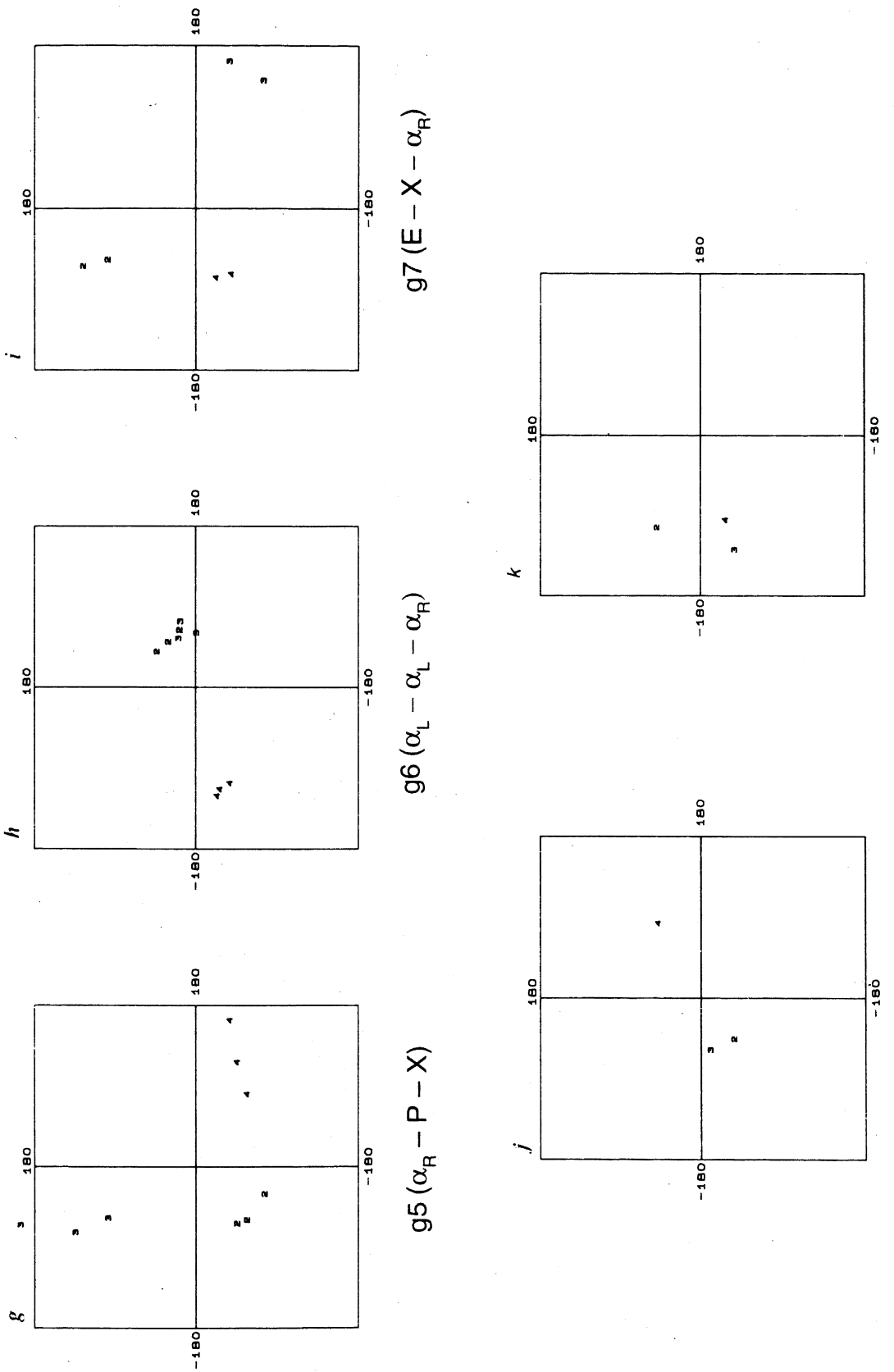


Figure 3. *a–k*, The plot of (ϕ, ψ) at the three α-carbon atoms in the α-turns in proteins grouped into different categories, based on conformational similarity. The conformational points corresponding to C_2^α, C_3^α and C_4^α are marked as 2, 3 and 4, respectively. The designation of each category is also marked.

Figure 3 *g–k*.

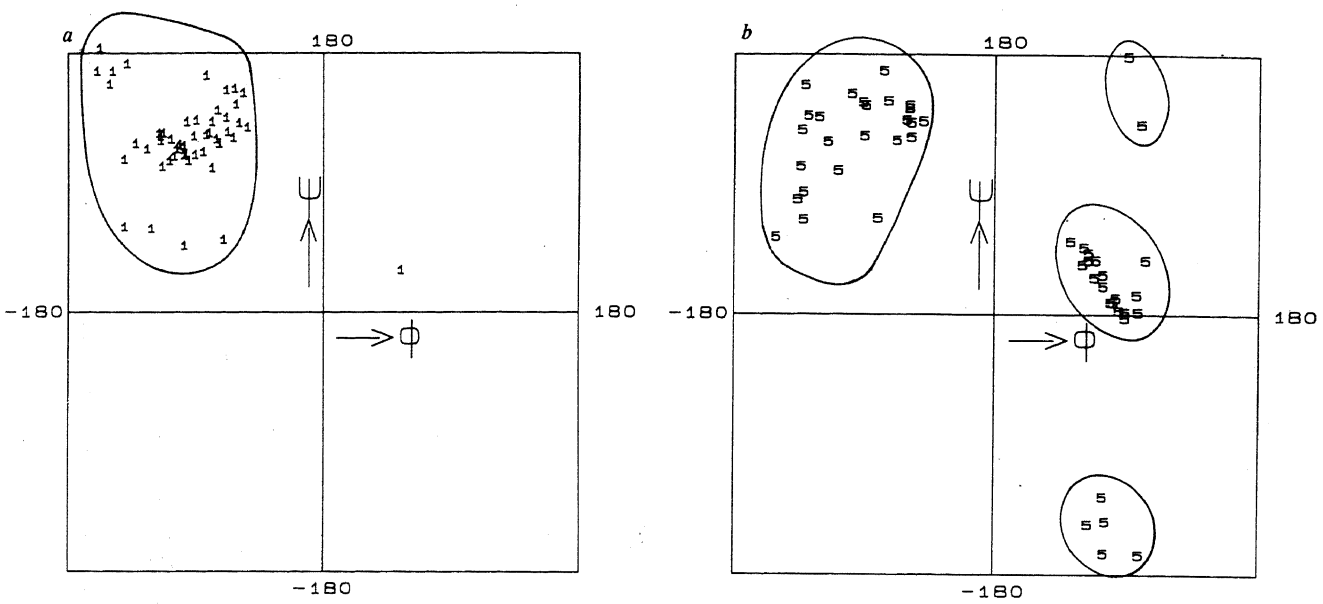


Figure 4. Plot of the (ϕ, ψ) values at the two α-carbon atoms immediately preceding and succeeding the α-turn in the examples of family F1. The clustering is also marked: *a*, at C_1^α; *b*, at C_5^α.

examples). It is very interesting to note that all the points except one in Figure 4 *a*, occur in the same quadrant of the (ϕ, ψ) plane, with a cluster near the extended conformation. There are two major and one minor clusters observable in Figure 4 *b*. One of the two major clusters occurs close to the extended region and the other in the α_L region. The minor cluster occurs in the inverse extended region (similar to the one designated as E′ in Figure 2 *a* and all these are glycyl residues). Thus, it is possible to have three motifs (while the term family is used to denote the segment with similar values at (ϕ, ψ) C_2^α, C_3^α and C_4^α, the term motif is used to denote the segment with similar (ϕ, ψ) values at all the five positions C_1^α to C_5^α) for this family, which can be represented by ($\beta - \alpha_R - \alpha_R - \alpha_R - \beta$), ($\beta - \alpha_R - \alpha_R - \alpha_R - \alpha_L$), ($\beta - \alpha_R - \alpha_R - \alpha_R - E'$), and the isolated example, which occurs in the protein enolase[21] (segment 102–106) can be designated as ($\alpha_L - \alpha_R - \alpha_R - \alpha_R - E'$).

The next largest family F2 has only 11 examples and is formed by conformational combination of E, α_L and α_R. This and other groups show that different combinations of (ϕ, ψ) at the three α-carbon atoms do occur in proteins, resulting in the formation of $5 \rightarrow 1$ hydrogen bonds to stabilize the local nonhelical regions. The designations of the families/groups are given in Table 1 *b*. Stereo diagrams of one representative example from each of the families F1, F2, groups g1–g7 and the two isolated examples are shown in Figure 5 *a–k*.

In the sections that follow, the different characteristics and features of α-turns are examined in detail.

End-to-end distances

The end-to-end distance (C_1^α–C_5^α) in the examples of α-turns varies between 4.7 and 6.7 Å. The distribution among the various families/groups in the different ranges is given in Table 2. (There are two examples of α-turns occurring in the protein 2cyp[22] (segment 35–39 belonging to F1) and in the protein 1gox[23] (segment 345–349 belonging to g1), which have end-to-end distances lying outside this range of 4.7–6.7 Å. The end-to-end distance is 4.2 Å in the former and 7.0 Å in the latter.)

The major family F1 has examples throughout the range, indicating flexibility in the chain segment. Groups g2, g3, g4 and g6 have smaller ranges of end-to-end distances (< 5.7 Å) and these groups have served the purpose of bringing the two ends of the segments much closer. On the other hand, the three members of group g5 are characterized by distances in the higher end of the spectrum (5.7–6.7 Å). It is worthwhile to mention that the computation of the corresponding distance in the various α-helical segments occurring in the data set shows that 92% of the examples lie in the range 4.0–7.0 Å (the ideal value for α-helix being 6.2 Å).

Propensity of amino acids to occur in α-helical α-turn

As mentioned earlier, the family F1, which contains α-helical α-turns, comprises a majority of examples. Since this is the shortest helix that can be formed, it

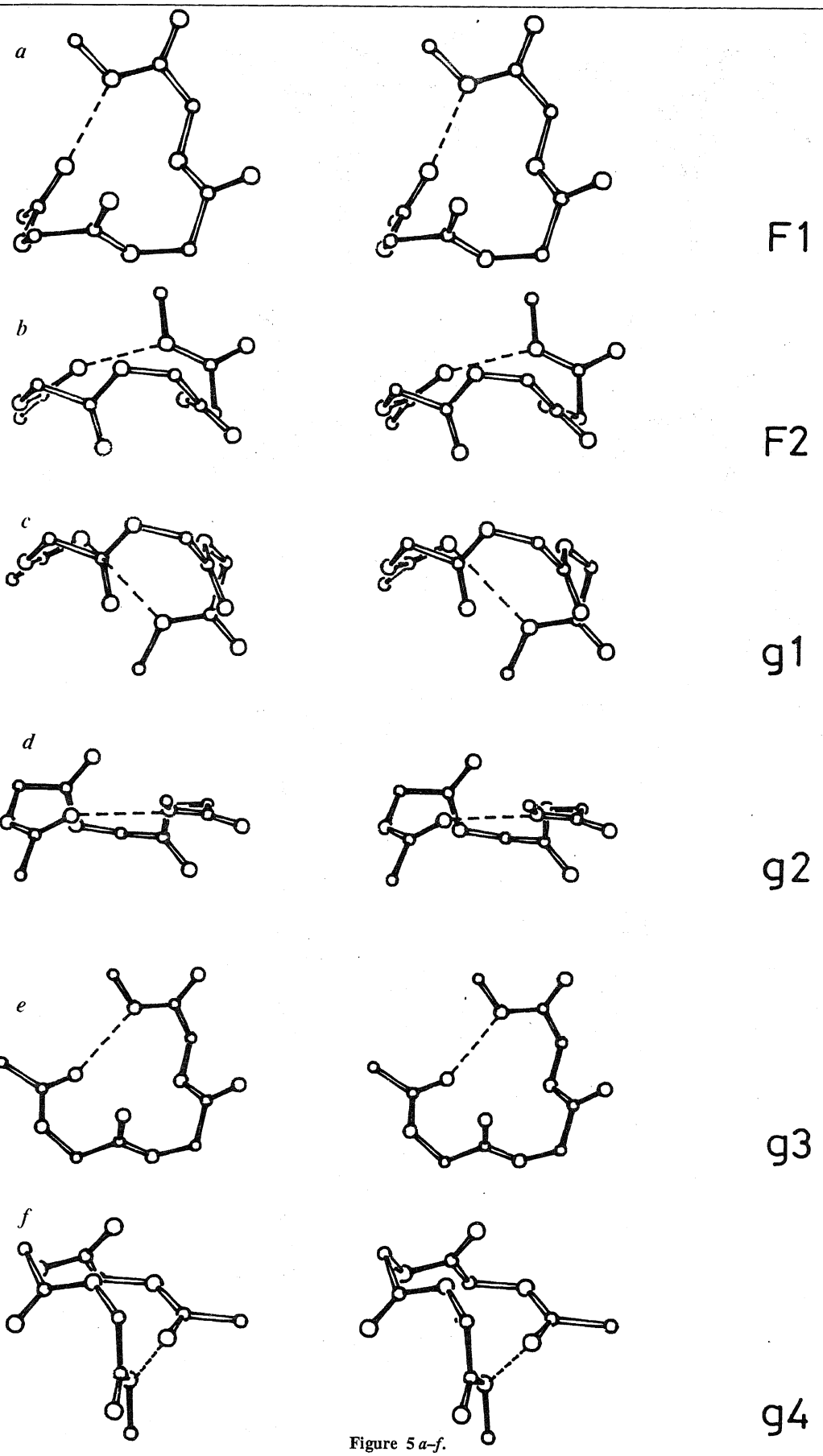


Figure 5 *a–f.*

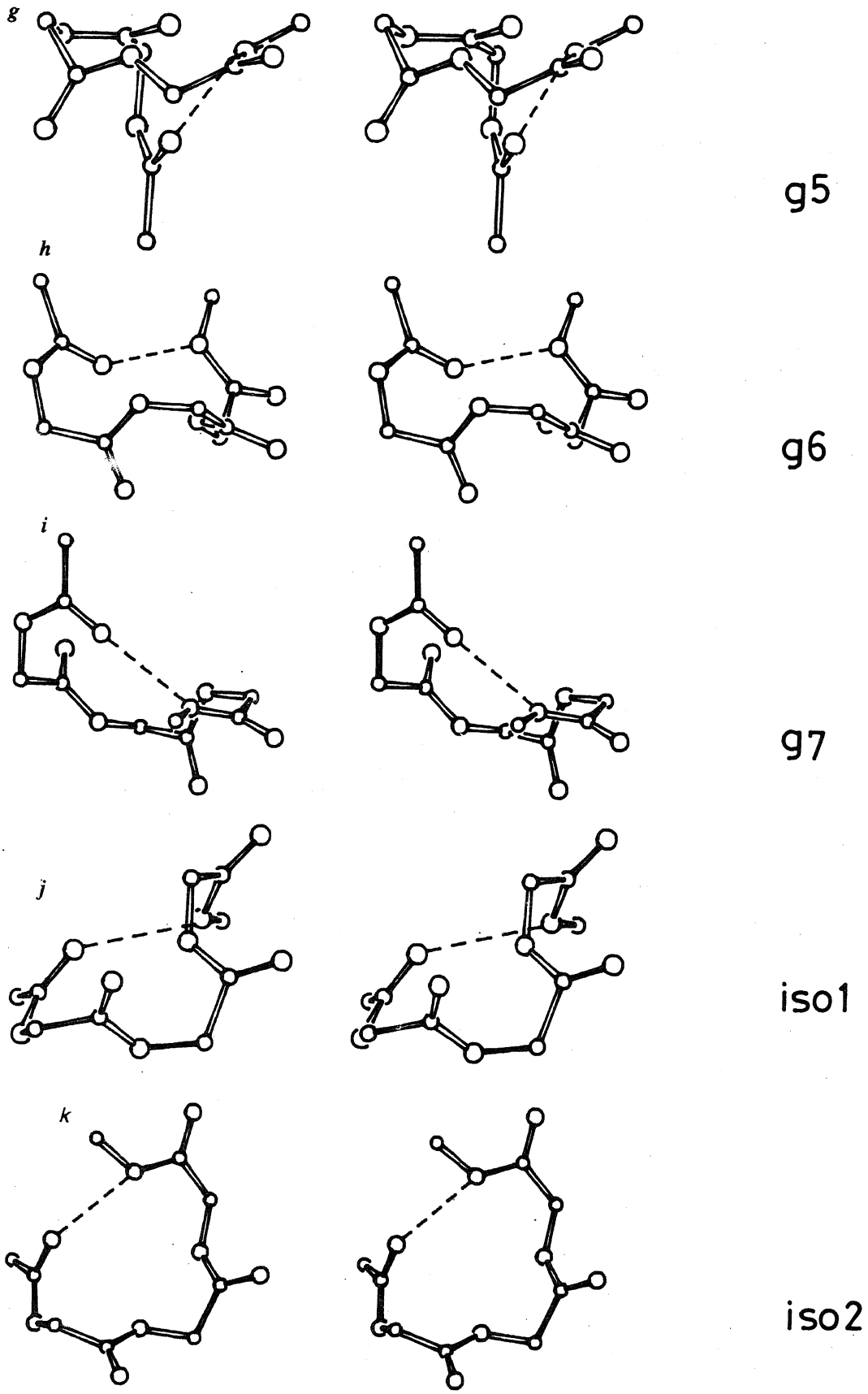


Figure 5 *g–k*.

Figure 5. *a–k*, ORTEP stereo diagrams of one representative example from different categories of α-turns. The 5 → 1 hydrogen bond characterizing the α-turn is shown as dotted lines and the category is marked. The diagrams have been taken using a version of ORTEP modified to suit PC.

Table 2. Distribution of end-to-end distances (C_1^α–C_5^α) in the examples of α-turns in different ranges

Family	Range (Å) 4.7–5.2	5.2–5.7	5.7–6.2	6.2–7.0	Total
F1	10	15	16	8	49
F2	2	6	1	—	9
g1	4	—	—	—	4
g2	1	1	—	—	2
g3	4	—	—	—	4
g4	1	3	—	—	4
g5	—	—	1	2	3
g6	2	1	—	—	3
g7	—	1	1	—	2
#	1	—	—	—	1
#	1	—	—	—	1
Total	26	27	19	10	82

– Isolated examples.

may be interesting to compare propensity values of amino acids involved in such turns. The propensity of occurrence of an amino acid at the specified position of the α-turn is given by

$$p_j^k = \frac{f_j^k}{F_j},$$

where f_j^k is the fraction of amino acid residue of type j at position k and F_j is the fraction of the same amino acid in the complete data set.

Using the above formula, the propensity of these amino acids to occur in the positions $i+1$, $i+2$, $i+3$ of the α-helical α-turn have been computed. These along with average propensity are given in columns 2–5 of Table 3. It is evident that four amino acids Cys, Glu, Ser and Thr show higher propensity (> 1.2) to occur in α-turns. (These are shown in bold in the table.)

Statistical significance of the propensity values has been determined using the *d*-test[11]. The statistic d is given by

$$d = \frac{(n \times f_j^k) - (n \times F_j)}{\sqrt{n \times F_j \times (1 - F_j)}},$$

where n is the number of examples in F1.

The amino acids with statistically significant propensity values for 5% tail ($d > 1.97$) are shown with an asterisk in Table 3. Of the four amino acids with high propensity values, only three (Glu, Ser and Thr) are seen to be statistically significant. The amino acid Cys, in spite of showing a high propensity value, is not statistically significant (probably due to the small sample size).

Since the α-turn when repeated, leads to an α-helix, it will be interesting to compare these propensity values with the corresponding ones for α-helical segments in proteins. The propensity values for amino acids to occur in an α-helix as computed using the data set used are

Table 3. Propensity of the 20 amino acids to occur in α-helical α-turn (family F1) and in α-helix. Those whose average propensity values are > 1.2 are in bold

Amino acid	Propensity to occur in α-turn in position $i+1$	$i+2$	$i+3$	Ave	α-helix
ALA	1.38*	1.38*	0.69	1.15	**1.48**
ARG	1.05	1.58	0.00	0.88	**1.23**
ASN	0.00	1.30	0.87	0.73	0.80
ASP	0.67	1.67*	1.00	1.11	0.94
CYS	**2.50**	**2.50**	**2.50**	**2.50**	0.78
GLN	0.57	0.57	1.71	0.95	**1.28**
GLU	1.54*	**2.69**	**2.31***	**2.18**	**1.46**
GLY	0.00	0.53	0.26	0.26	0.41
HIS	0.69	1.38	1.38	1.15	1.03
ILE	0.40	0.00	0.80	0.40	1.02
LEU	0.55	0.55	0.82	0.64	**1.31**
LYS	1.23*	1.54*	0.62	1.13	**1.27**
MET	0.80	0.00	0.00	0.27	**1.33**
PHE	0.00	0.00	1.11	0.37	1.06
PRO	**3.26***	0.00	0.00	1.09	0.56
SER	**2.19***	1.37*	0.82	**1.46**	0.75
THR	1.93*	0.97	**3.87***	**2.26**	0.75
TRP	0.00	0.00	0.83	0.28	1.14
TYR	0.59	0.00	0.00	0.20	0.85
VAL	0.63	1.25*	0.94	0.94	0.90

*Statistically significant as per *d*-test.

given in column 6 of Table 3. The values are in good agreement with those reported earlier by other workers[24,25]. The amino acids that exhibit high propensity (> 1.2) are shown in bold. (Glu, Met, Gln, Lys and Arg). It is interesting to note that, except for Glu, none of the other amino acids show large propensity both for α-helix and α-turn. This shows that the *amino acids that have larger propensity to occur in α-turn have no greater tendency to occur in α-helices and vice versa.* This qualitatively lends an explanation to the isolated nature of the α-turns. The above analysis shows that among the 20 amino acids, Ser and Thr have formed α-turns in spite of their low propensity for α-helices.

Another way of looking at the α-turns is to consider them as an overlap of two successive β-turns ($i \rightarrow i+3$ and $i+1 \rightarrow i+4$), with the β-turns being of type I and/or type III. Identifying β-turns by the criteria of Ramakrishnan and Srinivasan[26] (and not necessarily by the existence of $4 \rightarrow 1$ type hydrogen bond), 42 out of 50 examples of family F1 fall under this category (in 27 examples two type I β-turns overlap and in 15 examples type III and type I β-turns successively overlap). The propensities of these amino acids to occur in various types of β-turns have been reported by Wilmot and Thornton[11]. The interesting aspect is that the propensities of the amino acids Cys, Ser and Thr (which show greater propensity for α-turns) are also greater than 1.0 for type I β-turns. A further analysis shows that less than one-third of the overlapping β-turns (42 out of 153) alone form the α-turn.

Since all the 20 amino acids have occurred in α-turns, it would be more appropriate to look at the product propensity of the triplet of amino acids that occur at C_2^α, C_3^α and C_4^α (the product propensity being the product of the individual propensities) and select those triplets that may have a better preference to occur in α-turns as compared to α-helices. This is done by evaluating a quantity ΔP, which is the difference between the product propensity for an α-turn and the product propensity for α-helix. The distribution of ΔP is shown as a histogram in Figure 6. It can be seen that, though the peak occurs around 0, there are few triplets that occur in α-turns which have a large ΔP. Those with $\Delta P > 4.0$ are listed in Table 4. In the six distinct triplets listed in Table 4, Thr, Cys and Ser contribute 78%, which is to be expected since these are the predominantly high-propensity amino acids for α-turns.

Other hydrogen bonds of residues involved in α-turn segment

Since the isolated α-turns do not have consecutive hydrogen bonds to stabilize the segment, it may be interesting to find out whether these α-turn segments are involved in hydrogen bonds with other parts of the chain. On doing this, it was found that except in two examples, the segment is invariably involved with other hydrogen bonds (in addition to the $5 \rightarrow 1$ hydrogen bond). In fact, of the 252 residues in the α-turns (84 examples with three residues in positions $i+1$, $i+2$ and $i+3$), 178 (71%) act either as a donor or an acceptor for hydrogen bond. It is interesting to note that 79 out of 84 (94%) residues in position $i+3$ take part in hydrogen bonding. (The corresponding values for residues in positions $i+1$ and $i+2$ are 61% and 57%, respectively).

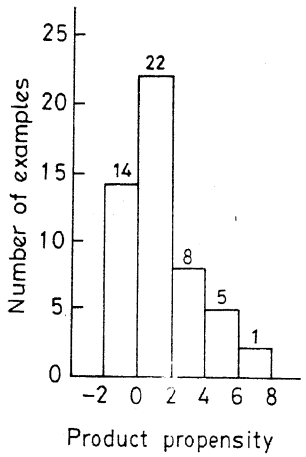


Figure 6. Histogram showing the distribution of the difference product propensity (ΔP) in the examples of α-turns in family F1 (see text for the definition of ΔP).

Table 4. Triplets having large preference to occur in α-turns arranged in decreasing order

Triplet	No. of examples	ΔP*
SER–THR–THR	1	7.027
CYS–ALA–THR	1	5.620
THR–HIS–THR	1	5.284
SER–CYS–SER	1	4.898
THR–VAL–CYS	1	4.769
CYS–ARG–THR	1	4.231

*Cut-off is 4.0.

Hydrophobicity of residues in α-turns

Another aspect that can be looked into is the nature (hydrophobic or otherwise) of the residues involved in the α-turn examples of family F1. Using the consensus values of Eisenberg[27], the average hydrophobicity of the α-turn segments (H_α^t) are evaluated (i.e. average of the hydrohobicity values of the residues at positions $i+1$, $i+2$ and $i+3$). This distribution is shown in Figure 7 *a*. Figure 7 *b* gives the corresponding distribution of hydrophobicity using the residues occurring in α-helices in the data set. It is clear that the residues involved in the isolated α-turns are more polar (hydrophilic) compared to those in α-helices. This would mean that these α-turns are more likely to occur on the surface, accessible to the solvent. This is in agreement with the now well-known concept[2] that the turns in protein usually tend to be on the surface as compared to the other longer secondary structures such as helices and sheets.

Location of α-turns in proteins

The α-turn, being a small segment with a well-defined geometry, can act as a linker between two secondary structural segments in proteins. This aspect has been looked into by examining the proximal secondary structure that (giving an allowance of maximum of four residues on either side) occurs on either side of the α-turn. The secondary structures were identified by examining the contiguous occurrence of representative (ϕ, ψ) values[26]. The results are tabulated in Table 5, in which H stands for helix, E for extended strand. (The commonly used notation is β-strands.) In an earlier communication[28] it has been shown that some of the β-strands are not involved in the formation of the β-sheet. The notation β_s is used to designate those which are involved in forming the β-sheet and β_i those which did not. We have preferred to use the notation E-strands since (i) these have been identified with the criteria that the (ϕ, ψ) values in the segment all lie in the

'extended conformational region' of Ramachandran map and (ii) no distinction has been made between β_i and β_s strands and U (uncharacterized) for other segments.

It can be seen from the table that there are examples for all the possible combinations. The maximum number of examples (19) occur for EE linkers (E–α_T–E). In fact, in about 70% of the examples the α-turn either precedes or succeeds the E-strand. This points to the tendency of α-turns to occur at the ends of extended strands.

The EE linkers are further examined from the point of view of immediate local chain reversal. Very interestingly, in 17 of these 19 examples, the α-turn occurs at the loop regions of β-hairpin. Even in the two cases 1sn3 (α-turn segment 7–11) and 2rsp (α-turn segment A46–50), the E-strands adjoining the α-turn though not forming a sheet are antiparallel in orientation and hence the full stretch can be considered as a hairpin bend without intermolecular hydrogen bonds. These are listed in Table 6, along with the α-turn families/groups to which they belong. Nine of these belong to family F1 and the remaining are distributed among the various groups. The β-hairpins occurring in proteins have been grouped into four classes by different workers[29–31]. Following the definition given by Pavone[30] the classes to which these linkers belong to have also been identified and these are also given in Table 6. Examples are found belonging to all the four classes (I, II, III and IV), with class IV having the maximum number (11).

Four examples, one in each from the families F1 (1cox)[32], g1 (1gcr)[33], g4 (3apr)[34] and one of the isolated examples (2fbj)[35] are shown in Figure 8, which brings out this aspect pictorially.

The α-turns which form linkers between other secondary structures (viz. (i) helix and helix, (ii) helix and strand and (iii) strand and helix) do not exhibit any conspicuous pattern or uniformity.

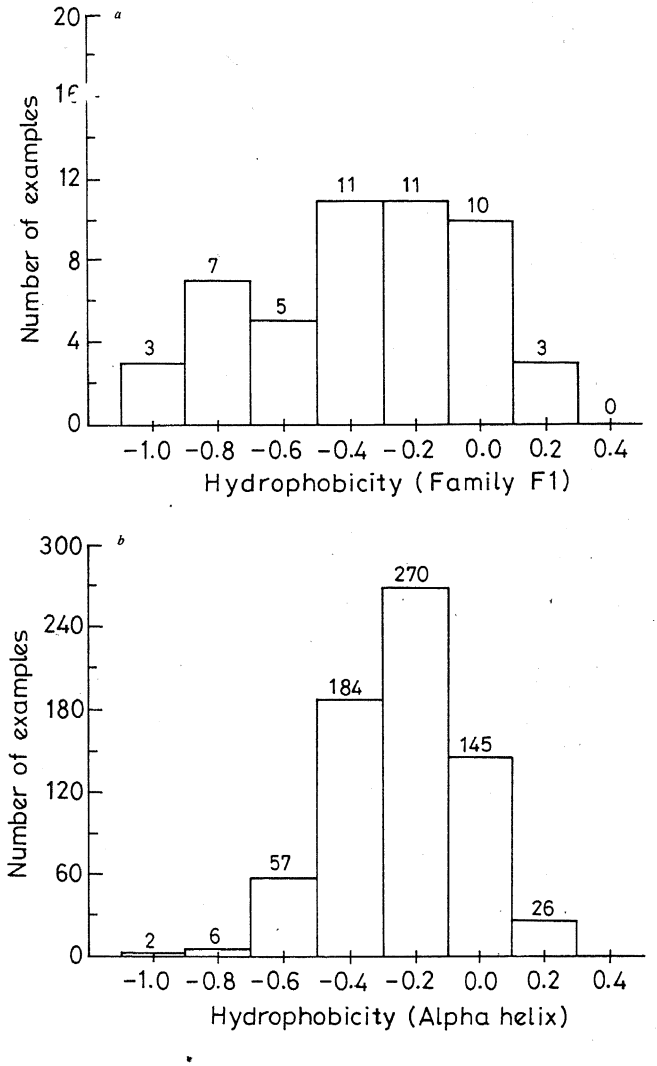


Figure 7. The distribution of average hydrophobicity: *a*, of α-turn segments in family F1; *b*, of α-helical segments in the data set.

Table 5. Number of examples of α-turns connecting different secondary structures (H–helix; E – extended strand and U – uncharacterized)

Sec. struc. preceding α-turn	Sec. struc. succeeding α-turn			
	H	E	U	Total
H	8	7	6	21
E	4	19	17	40
U	4	10	9	23
Total	16	36	32	84

Table 6. EE linker examples of α-turns which fall in different β-hairpin classes

Protein code	α-turn segment	Family/ group	β-hairpin class
1aap	A24–28	F1	IV
1cox	391–395	F1	IV
2er7	E240–244	F1	IV
2ltn	A54–58	F1	IV
2ltn	A167–171	F1	IV
2rhe	93–97	F1	IV
3blm	50–54	F1	IV
3cla	97–101	F1	IV
3fgf	67–71	F1	IV
3apr	E11–15	g4	IV
1ak3	A129–133	g7	IV
1gdl	O300–304	g6	III
4pep	9–13	g6	III
1rbp	63–67	#	III
1gcr	47–51	g1	II
1gcr	136–140	g1	II
2fbj	L166–170	#	I

– Isolated examples.

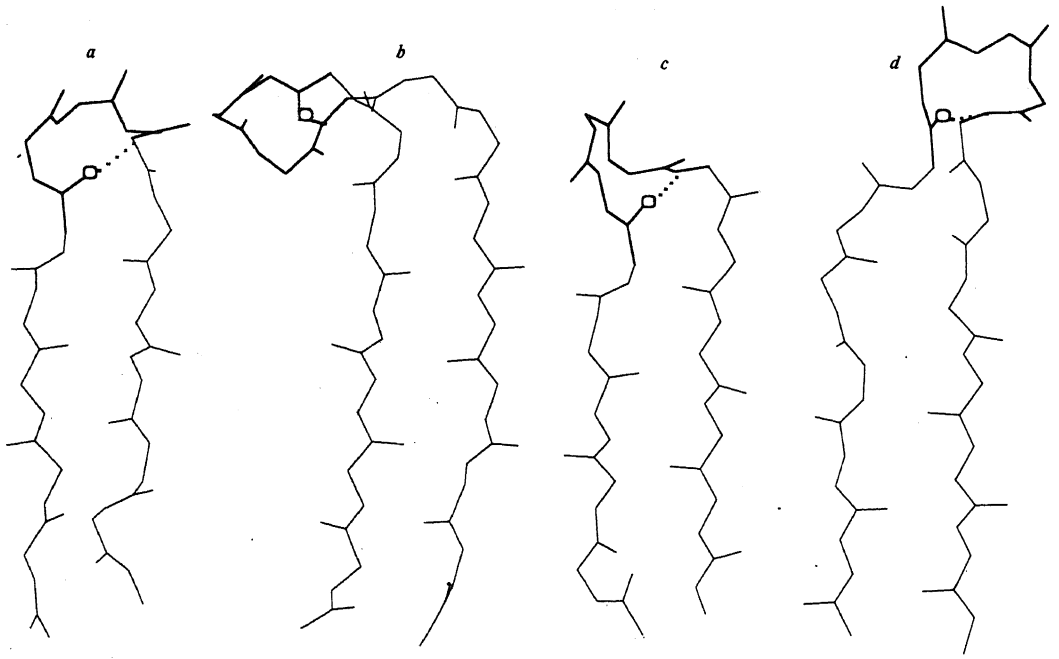


Figure 8. Two antiparallel β-strands connected by an α-turn. The α-turn segment is shown as thick lines and the 5 → 1 hydrogen bond by a dotted line. The oxygen atom taking part in the hydrogen bond is explicitly shown. *a*, segment 388–399 of the protein 1cox (α-turn 391–395: family F1); *b*, segment 41–58 of the protein 1gcr (α-turn 47–51: group g1); *c*, segment E5–E21 of the protein 3apr (α-turn E11–E15: group g4); *d*, segment L160–L172 of the protein 2fbj (α-turn L166–L170: isolated example).

Conclusion

The present study shows that the proteins do contain 5 → 1 hydrogen-bonded α-turns which do not necessarily form part of a larger secondary structure such as α-helix. The fact that the number of such examples of isolated α-turns is small as compared to β-turns, points to the fact that in these examples, α-turns have, for some reason, been preferred over the more populous β-turns. The small but significant number of examples of α-helical α-turns indicate that, in these cases, (i) either the α-helix which has been initiated has not progressed or, alternatively, (ii) the α-helix thus formed by the α-helical α-turn has not been able to propagate itself in either direction.

Since these examples occur in different types of proteins, the reason for these short α-helical α-turns may not lie so much in the functional/folding aspects of the protein as in the combination of residues in the α-turns as well as those which immediately precede and succeed these turns. In any case the isolated α-turns deserve to be considered as a secondary structural feature, though of a minor nature.

1. Richardson, J. S., *Adv. Protein Chem.*, 1981, **34**, 167–339.
2. Rose, G. D., Gierasch, L. M. and Smith, J. A., *Adv. Protein Chem.*, 1985, **37**, 1–109.
3. Richardson, J. S. and Richardson, D. C., *Trends Biochem. Sci.*, 1989, **14**, 304–309.
4. Nemethy, G. and Printz, M. P., *Macromolecules*, 1972, **5**, 755–758.
5. Matthews, B. W., *Macromolecules*, 1972, **5**, 818–819.
6. Venkatachalam, C. M., *Biopolymers*, 1968, **6**, 1425–1436.
7. Chou, P. Y. and Faman, G. D., *J. Mol. Biol.*, 1977, **115**, 135–175.
8. Efimov, A. V., *Molecular Biology (USSR)*, 1986, **20**, 258–266.
9. Thornton, J. M., Sibanda, B. L., Edwards, M. S. and Barlow, D. J., *BioEssays*, 1988, **8**, 63–69.
10. Srinivasan, N., Sowdhamini, R., Ramakrishnan, C. and Balaram, P., in *Molecular Conformation and Biological Interactions* (eds Balaram, P. and Ramaseshan, S.), Indian Academy of Sciences, Bangalore, 1991.
11. Wilmot, C. M. and Thornton, J. M., *J. Mol. Biol.*, 1988, **203**, 221–232.
12. Wilmot, C. M. and Thornton, J. M., *Prot. Eng.*, 1990, **3**, 479–493.
13. Milner-White, E. J., Ross, B. M., Ismail, R., Bethadj-Mostefa, K. and Poet, R., *J. Mol. Biol.*, 1988, **204**, 777–782.
14. Milner-White, E. J., *J. Mol. Biol.*, 1990, **216**, 385–397.
15. Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Brice, Jr, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. and Tasumi, M., *J. Mol. Biol.*, 1977, **112**, 535–542.
16. Abola, E. E., Bernstein, F. C., Bryant, S. H., Koetzle, T. F. and Weng, J., in *Crystallographic Databases – Information Content, Software Systems, Scientific Applications*, (eds Allen, F. H., Bergerhoff, G. and Sievers, R.), Data Commission of International Union of Crystallography, Bonn, 1987, pp. 107–132.
17. Mitra, J. and Ramakrishnan, C., *Int. J. Pept. Protein Res.*, 1981, **17**, 401–411.
18. Corey, R. B. and Pauling, L., *Proc. R. Soc. London*, 1953, **B141**, 10–20.
19. Stein, P. E., Leslie, A. G. W., Finch, J. T. and Carrell, R. W., *J. Mol. Biol.*, 1991, **221**, 941–959.
20. Ramachandran, G. N. and Sasisekharan, V., *Adv. Protein Chem.*, 1968, **23**, 283–437.
21. Lebioda, L. and Stec, B., *J. Am. Chem. Soc.*, 1989, **111**, 8511–8513.

22. Finzel, B. C., Poulos, T. L. and Kraut, J., *J. Biol. Chem.*, 1984, **259**, 13027–13036.
23. Lindqvist, Y., *J. Mol. Biol.*, 1989, **209**, 151–166.
24. Chou, P. Y. and Fasman, G. D., *Biochemistry*, 1974, **13**, 211–222.
25. Leszczynski, J. F. and Rose, G. D., *Science*, 1986, **234**, 849–855.
26. Ramakrishnan, C. and Srinivasan, N., *Curr. Sci.*, 1990, **59**, 837–847.
27. Eisenberg, D., Weiss, R. M., Terwilliger, T. C. and Wilcox, W., *Faraday Symp. Chem. Soc.*, 1982, **17**, 109–120.
28. Soman, K. V. and Ramakrishnan, C., *Int. J. Biol. Macromol.*, 1986, **8**, 89–96.
29. Milner-White, E. J., and Poet, R., *Biochem. J.*, 1986, **240**, 289–292.
30. Pavone, V., *Int. J. Biol. Macromol.*, 1988, **10**, 238–240.
31. Sibanda, B. L., Blundell, T. C. and Thornton, J. M., *J. Mol. Biol.*, 1989, **206**, 759–777.
32. Vrielink, A., Lloyd, L. F. and Blow, D. M., *J. Mol. Biol.*, 1991, **219**, 533–554.
33. Summers, L., Wistow, G., Narebor, M., Moss, D., Lindley, P., Slingsby, C., Blundell, T., Bartunik, H. and Bartels, K., *Pept. Protein Rev.*, 1984, **3**, 147–168.
34. Suguna, K., Padlan, E. A., Smith, C. W., Carlson, W. D. and Davies, D. R., *Proc. Natl. Acad. Sci. USA*, 1987, **84**, 7009–7013.
35. Bhat, T. N., Padlan, E. A. and Davies, D. R., (private communication through PDB).

ACKNOWLEDGEMENTS. We thank Prof. P. Balaram and Dr N. V. Joshi for their suggestions and comments.

Received 24 July 1995; accepted 31 July 1995

MEETINGS/SYMPOSIA/SEMINARS

Winter School on 'NMR in Structural Biology' (an event for the Golden Jubilee Celebrations of TIFR)

Date: 5–18 November 1995
Place: TIFR, Bombay

Topics include: Multidimensional high resolution NMR spectroscopy in liquids; Introduction to pulse/field gradients; Applications to nucleic acids and proteins; Biomembranes/cellular NMR; NMR imaging; Molecular dynamics/molecular simulations.

Contact: Dr (Mrs) Sudha Srivastava
National Facility for High Field NMR
Tata Institute of Fundamental Research
Homi Bhabha Road
Bombay 400 005, India
Phone: 22-215 2971 Ext. 2271
Telex: 011-83009 TIFR IN
Fax: 091-22-2152110
e-mail: sudha@tifrvax.tifr.res.in

Sixth All India Conference on Cytology and Genetics

Date: 23–25 February 1996
Place: Rohtak

Contact: Dr M. S. Chennaveeraiah
Hon. Secretary
Society of Cytologists and Geneticists
No. 9, Byrasandra Main Road
First Block East, Jayanagar
Bangalore 560 011, India

International Conference on Natural and Technological Coastal Hazards

Date: 2–6 December 1995
Place: Tirupati, India

Contact: Prof. K. L. Narasimha Rao
Department of Geology
Sri Venkateswara University
Tirupati 517 502, India
Fax: +91-8574-24111/25211

Second Congress on Traditional Sciences and Technologies of India

Date: 27–31 December 1995
Place: Anna University, Madras

Subjects include: Agriculture and livestock management; Forestry, water management, fisheries, bio-diversity; Health-care, food and nutrition; Architecture and building materials; Metallurgy and metal-working; Textiles and natural dyes; Pottery, leather, wood and bamboo, other industries; Ship-building and navigational technologies; Indian mathematics.

Themes include: Role of communities; Education and training; Local markets; Role of women; Economic policies for traditional S&T.

Contact: The Madras Secretariat
Traditional S&T Congress
Students' Centre, Anna University
Madras 600 025, India
Phone: (044) 2351126-Extn: 3113
Fax: (044) 2368403/2350397

Task activation of the human brain studied in real time by functional magnetic resonance imaging

N. R. Jagannathan* and P. Raghunathan

Department of Nuclear Magnetic Resonance, All India Institute of Medical Sciences, Ansari Nagar, New Delhi 110 029, India

This paper describes our initial results on functional brain mapping using a whole-body clinical magnetic resonance (MR) scanner operating at 1.5 T and having no echo planar imaging capability. The focus is on the effects of motor cortex stimulation of normal volunteers using a circularly polarized head coil when a gradient echo pulse sequence is employed. The study demonstrates the technical feasibility of functional MR imaging by using conventional gradient echo sequence and equipment.

MAGNETIC resonance imaging (MRI) is a well-established noninvasive clinical diagnostic tool which produces a spatial display of hydrogens present in the human body and provides a high-resolution morphological picture of the anatomy with superior contrast resolution[1]. Recent developments in MRI methodology have demonstrated the possibility of observing human brain function[2–15]. 'Functional' imaging (fMRI) has great potential for assessing individual pathophysiology and for characterization of human brain functions such as language, learning, memory, task activation, etc.[16]. Using the echo planar imaging (EPI) technique, Belliveau *et al.*[2] described the measurement of blood volume during visual activation, which not only opened up exciting new possibilities in the study of brain physiology but also provided a unique and challenging tool for neuroscience. Since then, several research groups have demonstrated that functional imaging of the human brain was feasible on their conventional MR system, where EPI capabilities were not available[6–15].

In this paper, we present our initial experimental results on task activation studies using the fMRI technique. Our studies have been carried out on normal volunteers using a Siemens MAGNETOM 63SP-4000 whole-body scanner. Our objective was to evaluate and demonstrate the feasibility of such studies on our 1.5 T whole-body scanner, using the FLASH sequence[17].

Experiments were performed on the 1.5 T superconducting whole-body scanner using a circularly polarized head coil. Magnetic field shimming with all first-order coils was performed by means of global shimming for each subject, achieving typical line widths of about 15 Hz. Scout images, one in each orthogonal plane, were initially obtained. Later, multislice T_1-weighted images in the sagittal, and in some cases coronal, planes were obtained using the standard spin echo sequence with an echo time (TE) of 15 ms, repetition time (TR) of 520 ms, a 256×256 matrix and a slice thickness of 5 mm. Using these images as an anatomic guide, appropriate tilted planes were then selected for task activation studies. One or more of the tilted axial planes passing through the primary motor and sensory cortex (Figure 1) were selected. The pulse sequence used was a conventional two-dimensional, refocused gradient echo FLASH sequence[17,18] with TR = 90 ms, TE = 60 ms, flip angle = 40°, matrix size 64×128, field of view (FOV) = 200 mm, and slice thickness = 8 mm.

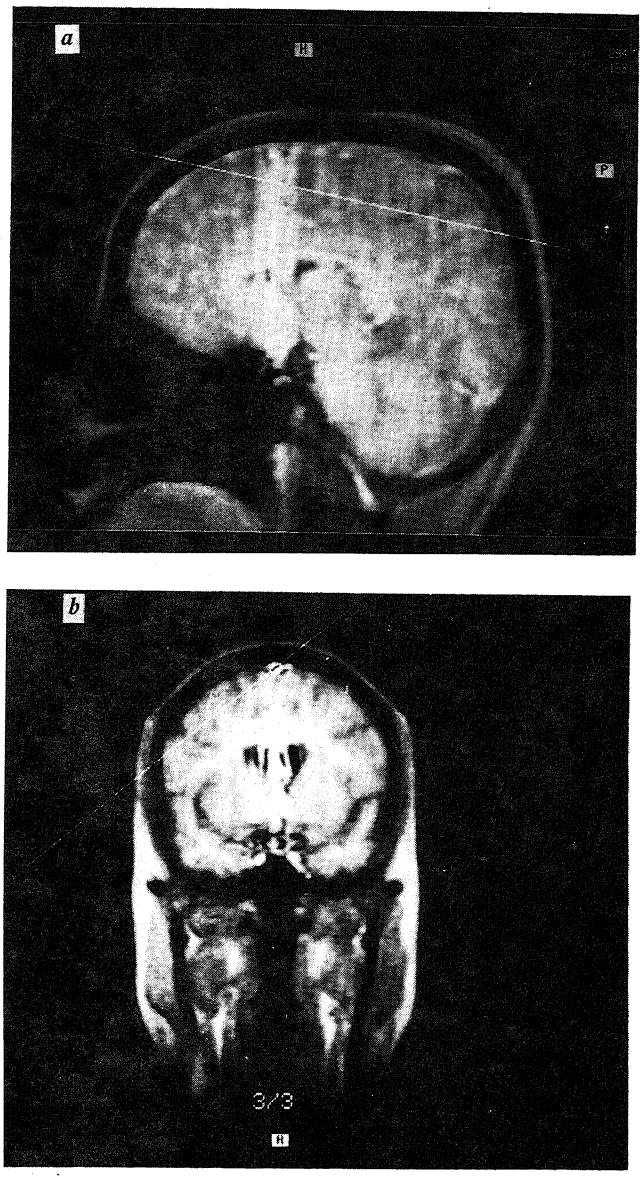


Figure 1. Proton MR images illustrating the orientation of the axial planes selected for task activation studies: *a*, for observing both the hemispheres of the brain simultaneously; *b*, for observing the left-hand movement on the right hemisphere of the brain.

*For correspondence.

These were optimized parameters, selected on the basis of a number of preliminary experiments. The percentage increase in signal intensity varies approximately linearly with TE; however, due to T_2^* processes, the overall signal intensity in both the control and activated images are known to decline with TE for a chosen tip angle of the RF pulse[9]. The maximum signal intensity increase in absolute terms occurs at a TE value approximately equal to T_2^*, which is around 66 ms in the activated region at 1.5 T (refs 6, 9). Our preliminary experiments also gave similar results and hence a TE value of 60 ms was used for all fMRI studies.

The hand task activation consisted of the volunteers ($n = 4$) either holding a sponge in the left or right hand and squeezing it repetitively, or performing finger-to-thumb movements. The rate and force of the squeezing of the sponge and finger-to-thumb movements were not controlled. Data were collected in blocks of 10 control images (i.e. in the task-free resting state), followed by 10 obtained during task activation (Figure 2). This procedure was repeated 4 or 5 times for each hand movement. When the axial plane shown in Figure 1 *a* was used, both left and right motor cortices were observed simultaneously. For this study, the volunteers ($n = 3$) held the sponges in both the hands and were asked to squeeze them simultaneously.

The MR images were reconstructed in the standard manner. Necessary postprocessing such as addition and subtraction of images was carried out using specialized software packages.

Figure 3 *a* shows an oblique T_1-weighted spin echo image of the anatomy of the selected section for subsequent monitoring of task activation, obtained using Figure 1 *b* as a pilot image. Figure 3 *b* shows the effects of left-hand movement; the appropriate cortical activity can be clearly seen near the sulcal wall. For observing both left and right motor cortices simultaneously, an axial oblique through a sagittal section was used as shown in Figure 1 *a*. Figure 3 *c* shows one such activated image when the volunteer was asked to activate both the hands, as described earlier. In general, the maximal signal intensity increase for the appropriate slice thickness used has been computed to be around 3–8%.

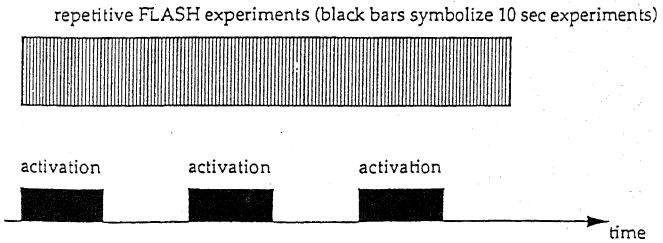


Figure 2. Experimental protocol for motor cortex activation. During 'activation' the volunteer performs the task of squeezing the sponge or finger-to-thumb motion.

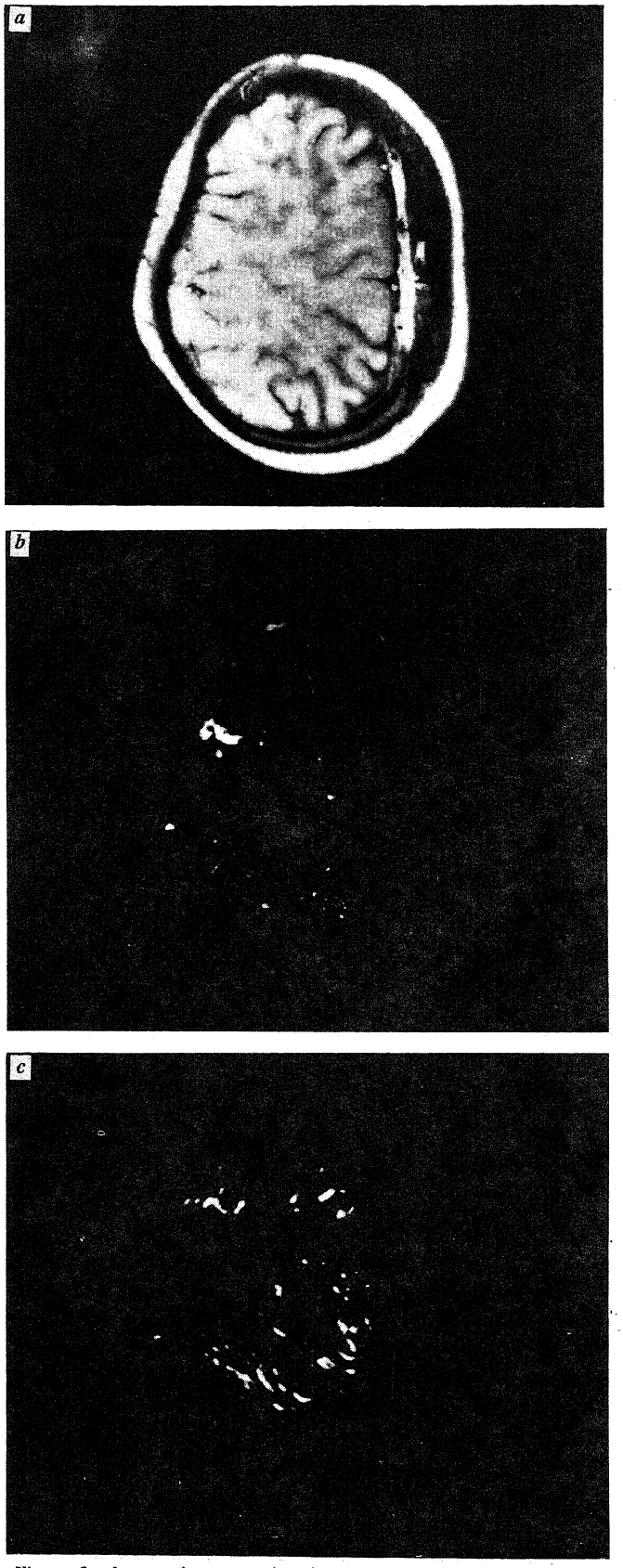


Figure 3. Images from a study of task activation obtained using the section orientation shown in Figure 1 *b*. *a*, Anatomy of the selected section. *b*, Effects of activation of the left-hand task. *c*, Effects of activation during the task performed using both the hands simultaneously.

The images shown in Figure 3 *c* are suitable for simultaneous examination of both left and right motor cortices. However, oblique planes through one hemisphere of the brain are preferable for viewing the detailed functional anatomy of the primary motor and sensory cortices. It is clear from Figures 3 *b* and *c* that activation is seen in grey matter of the sulcal walls, with little or no apparent effect in the intervening white matter[9]. In addition, the region activated by the hand movement task can also be precisely defined and related to specific gyri and sulci at a deeper level by following the activation down through successive sections using the three different planes shown in Figure 4 *a*. The set of images thus obtained from the right hemisphere of the brain from the superficial and middle planes are shown in Figures 4 *b* and *c* for the left-hand movement (see figure caption for details). As is seen clearly, the activation manifests itself as a hyperintense patch in the grey matter on both walls of the central sulcus. Other minor signal enhancement regions are also observed, showing that the area of activation is not confined to the walls and to the central sulcal region. These enhancement regions are believed to be associated with venous microvasculature in the grey matter[3].

The activation experiments were performed by first acquiring a set of 10 'activated' images – series A (during activation) – followed by 10 corresponding control rest images – series B (without any hand movement). This procedure of 'activity' (series A) followed by 'rest' (series B) was repeated 3 or 4 times. The image processing algorithm consists of adding together all the images of set A and producing a composite 'ACTIVE' image A. Then all images of series B, corresponding to rest period (no activation), were added together to produce a composite 'REST' image B. Finally, the algorithm takes the difference (A–B), and displays it as the 'functional MRI' image C. In this manner, the activation images shown in Figures 3 and 4 are the difference images obtained after 3 or 4 cycles of rest and activity. The accumulation of each data set takes about 800 s. The increase in signal intensity varied between 3 and 8%, depending upon the slice thickness used. However, the percentage signal intensity change depends on a number of factors such as partial volume effects, T_2^*, and other parameters used, such as the echo time, repetition time, slice thickness and, of course, the field strength. Intensity variations have also been known to occur if the postprocessing procedure is not correct. Subject movement during activation experiments may correlate with the task being performed and signal changes noticed may then be mistaken for evidence of brain activation[13]. Recently, several reports have appeared in the literature on these aspects and methods have been suggested to overcome these[19,20].

In our work, proper care has been taken regarding

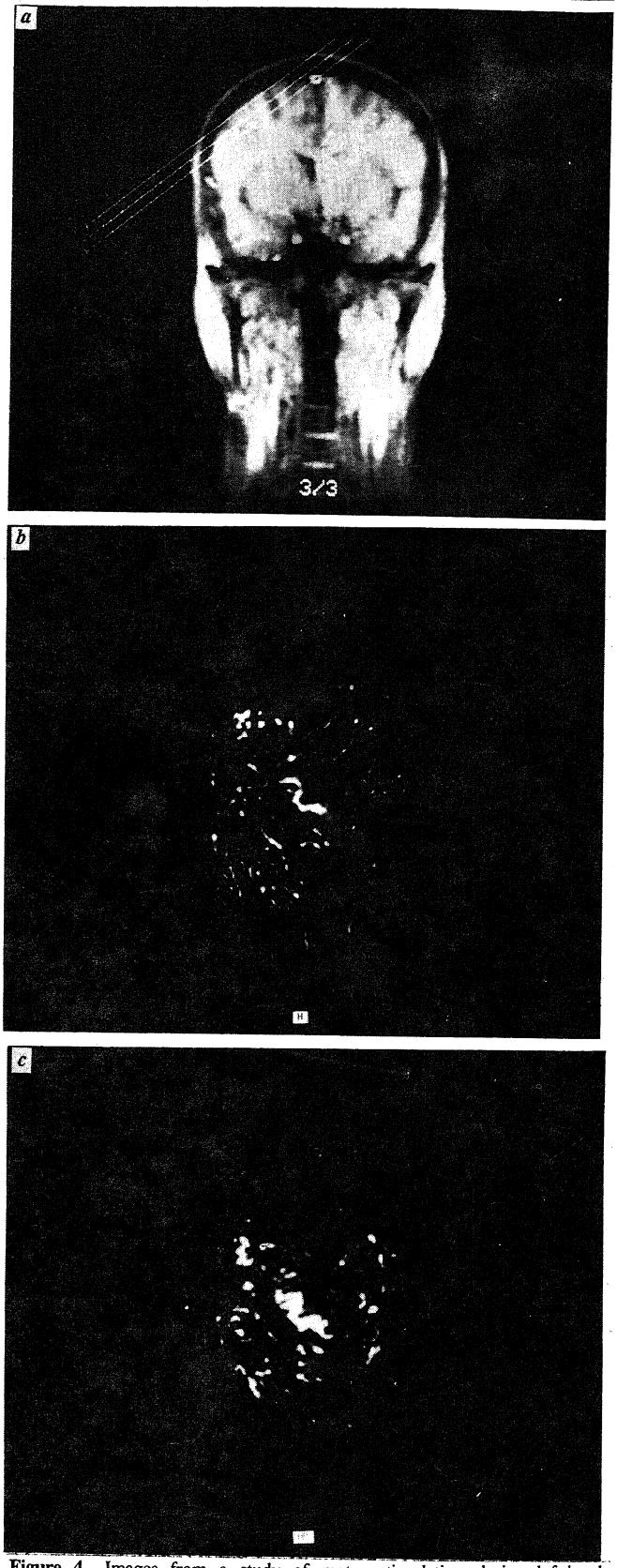


Figure 4. Images from a study of motor stimulation during left-hand movement task. *a*, The different contiguous planes selected; superficial, middle and deep. *b*, Activation in the superficial plane. *c*, Image from the middle plane. (The deep plane contains no image enhancement, denoting that it was not activated in hand movement and hence is not shown.)

subject movement during data acquisition to reduce motional artefacts. In addition, a set of 10 control (rest) images (with no activation – called set X) followed by another set of 10 control (rest) images (set Y) were separately acquired. All the images of set X were added together, producing a net 'REST' image X; similarly, all the images of set Y were added to produce another composite 'REST' image Y. The difference image, i.e. X–Y, was then obtained, which showed no signal enhancement in the plane selected, thus justifying that artefactual effects had been kept down to a minimum in executing our postprocessing algorithm.

The increase in local signal intensity observed in Figures 3 and 4 may be explained on the basis of our understanding of the blood-oxygenation-level-dependent (BOLD) mechanism whereby changes in neuronal activity affect the MR signal[21,22]. During task activation, brain stimulation takes place with increased blood flow, and oxygen delivery to the activated region exceeds the metabolic need[23]. The rise in blood supply seems to exceed the rise in oxygen consumption, leading to a local decrease in the ratio of deoxyhaemoglobin to oxyhaemoglobin[21,22]. Deoxyhaemoglobin is a paramagnetic substance and acts as an effective endogenous contrast agent. Moreover, a decrease in deoxyhaemoglobin concentration also decreases the vessel tissue susceptibility differential (leads to increased T_2^*), allowing increased spin coherence and thus increased signal while using a gradient echo sequence such as FLASH.

The functional MR images presented in this paper demonstrate that it is possible to use a conventional 1.5 T MR imager without echo planar capability for monitoring the brain functions if a gradient echo pulse sequence such as FLASH is employed. The application potential of fMRI is enormous in the area of neurosciences for assessing individual pathophysiology and for characterization of several brain functions[16] related to language[24], memory etc. In addition, it may also have important applications in neurosurgery. Identification of the anatomical relationship of a functional area to a tumour (which often distorts and displaces normal anatomy) is a great help to the neurosurgeon while planning the surgical approach and to preserve these primary areas during therapeutic operations[9].

1. Stark, D. and Bradley, W. (eds), *Magnetic Resonance Imaging*, Mosby, St. Louis, MO, 1992, 2nd edn.
2. Belliveau, J. W., Kennedy, D. N., Mckinstry, R. C., Buchbinder, B. R., Weisskoff, R. M., Cohen, M. S., Vevea, J. M., Brady, T. J. and Rosen, B. R., *Science*, 1991, **254**, 716–719.
3. Ogawa, S., Tank, D. W., Menon, R. S., Ellermann, J. M., Kim, S. G., Merkle, H. and Ugurbil, K., *Proc. Natl. Acad. Sci. USA*, 1992, **89**, 5951–5955.
4. Bandettini, P. A., Wong, E. C., Hinks, R. S., Tikofsky, R. S. and Hyde, J. S., *Magn. Reson. Med.*, 1992, **25**, 390–397.
5. Turner, R., Jezzard, P., Wen, H., Kwong, K. K., LeBihan, D., Zeffio, T. and Balaban, R., *Magn. Reson. Med.*, 1993, **29**, 277–279.
6. Balmire, A. M., Ogawa, S., Ugurbil, K., Rothman, D., McCarthy, G., Ellerman, J. M., Hyder, F., Rattner, Z. and Shulman, R. G., *Proc. Natl. Acad. Sci. USA*, 1992, **89**, 11069–11073.
7. Frahm, J., Bruhn, H., Merboldt, K. and Hanicke, W., *J. Magn. Reson. Imaging*, 1992, **2**, 501–505.
8. Frahm, J., Merboldt, K. and Hanicke, W., *Magn. Reson. Med.*, 1993, **29**, 139–144.
9. Connelly, A., Jackson, G. D., Fracknowiak, R. S. J., Belliveau, J. W., Vargha-Khadem, F. and Gadian, D. G., *Radiology*, 1993, **188**, 125–130.
10. Kim, S. G., Ashe, J., Hendrich, K., Ellermann, J. M., Merkle, H., Ugurbil, K. and Georgopoulos, A. P., *Science*, 1993, **261**, 615–617.
11. Schneider, W., Noll, D. C. and Cohen, J. D., *Nature*, 1993, **365**, 150–153.
12. Duyn, J. H., Moonen, C. T. W., deBoer, R. W., vanYperen, G. H. and Luyten, P. R., *NMR Biomed.* 1994, **7**, 83–88.
13. Constable, R. T., McCarthy, G., Allison, T., Anderson, A. W. and Gore, J. C., *Magn. Reson. Imaging*, 1993, **11**, 451–459.
14. Schad, L. R., Trost, U., Knopp, M. V., Muller, E. and Lorenz, W. J., *Magn. Reson. Imaging*, 1993, **11**, 461–464.
15. Constable, R. T., Kennan, R. P., Puce, A., McCarthy, G. and Gore, J. C., *Magn. Reson. Med.*, 1994, **31**, 686–690.
16. Rosen, B. R., Belliveau, J. W., Aronen, H. J., Hamberg, L. M., Kwong, K. K. and Fordham, J. A., in *Magnetic Resonance Neuroimaging* (eds Kucharczyk, J., Moselay, M. and Barkovich, A. J.), CRC Press, Boca Raton, 1994, pp. 141–166.
17. Hasse, A., Frahm, J., Marthaei, D., Hanicke, W. and Merboldt, K. D., *J. Magn. Reson.*, 1986, **67**, 258–266.
18. Hasse, A., Frahm, J., Matthaei, D., Merboldt, K. D. and Hanicke, W., in *Book of Abstracts*: Fourth Annual Meeting of the Society of Magnetic Resonance in Medicine, SMRM, London, 1985, pp. 980–981.
19. Hajnal, J. V., Myers, R., Oatridge, A., Schwieso, J. E., Young, I. R. and Bydder, G. M., *Magn. Reson. Med.*, 1994, **31**, 283–291.
20. Hu, X. and Kim, S. G., *Magn. Reson. Med.*, 1994, **31**, 495–503.
21. Ogawa, S., Lee, T. M., Kay, A. and Tank, D., *Proc. Natl. Acad. Sci. USA*, 1990, **87**, 9868–9872.
22. Ogawa, S., Lee, T. M., Nayak, A. S. and Glynn, P., *Magn. Reson. Med.*, 1990, **14**, 68–78.
23. Fox, P. T. and Raichle, M. E., *Proc. Natl. Acad. Sci. USA*, 1986, **83**, 1140–1144.
24. Shaywitz, B. A., Shaywitz, S. E., Pugh, K. R., Constable, R. T., Skudlarski, P., Fulbright, R. K., Bronen, R. A., Fletcher, J. M., Shankweiler, D. P., Katz, L. and Gore, J. C., *Nature*, 1995, **373**, 607–609.

ACKNOWLEDGEMENTS. We thank Profs P. N. Tandon and A. K. Banerji for their interest, encouragement and critical evaluation of the manuscript, Dr V. Govindaraju for many useful discussions and Dr R. Jayasundar for making available the processing software from Erlangen. Financial support from the Department of Science and Technology and Department of Biotechnology in the initial acquisition of MRI equipment is gratefully acknowledged.

Received 20 May 1995; revised accepted 10 July 1995

Radon precursory signals of Chamba earthquake

H. S. Virk, Vivek Walia and Anand K. Sharma

Department of Physics, Guru Nanak Dev University, Amritsar 143 005, India

Radon is recognized as a plausible precursor in earthquake prediction research. Radon monitoring in soil-gas and groundwater at Chamba and Dalhousie stations is carried out by using alpha-logger probes and emanometry. Radon anomalies are generally followed by seismic events in the grid (30–34°N, 74–78°E). The Chamba earthquake of 5.1 M, which occurred on 24 March 1995, had recorded its precursory signals.

Since the first observation of episodic changes in radon concentration in groundwater prior to the 1966 Tashkent earthquake[1], radon has been considered to be a plausible precursor for earthquake prediction. Observations of radon in soil-gas and groundwater have been recorded by various authors[2–10] and correlated with some of the earthquakes which occurred in USA, Japan, China, Mexico and India.

The physical basis of the radon anomalies prior to an impending earthquake have yet to be fully understood in terms of a comprehensive theoretical model. However, it is understood that the build-up of stresses prior to a seismic event induces strain fields in the crustal rocks; as a consequence, the radon emanation is enhanced. Model calculations indicate that a change in strain fields of at most 10^{-6}–10^{-8} caused the radon anomalies.

Radon monitoring work was started at Chamba and Dalhousie stations in Himachal Pradesh during March 1993 under the Himalayan Seismicity Programme of the Department of Science and Technology (DST), Government of India. Continuous radon monitoring is carried out in the soil-gas at Chamba and Dalhousie using the alpha-logger system, while radon is monitored daily at Dalhousie station in soil-gas and groundwater using the emanometry technique. The location of the monitoring stations is shown in the map (Figure 1), which shows major thrust faults in N–W Himalaya.

The alpha-logger probe (manufactured by Alpha-Nuclear Company in Toronto, Canada) is a portable battery-powered microprocessor-based data acquisition and control system. The detector probe consists of a silicon-diffused junction for the detection of alpha particles and can record radon alpha counts in 15 min increments over a period of 40 days non-stop. The detector unit is placed inside a covered auger hole about 60 cm in depth. The recorded data are retrieved with the aid of a laptop IBM-compatible PC by entering the proper commands. The software supplied with the system

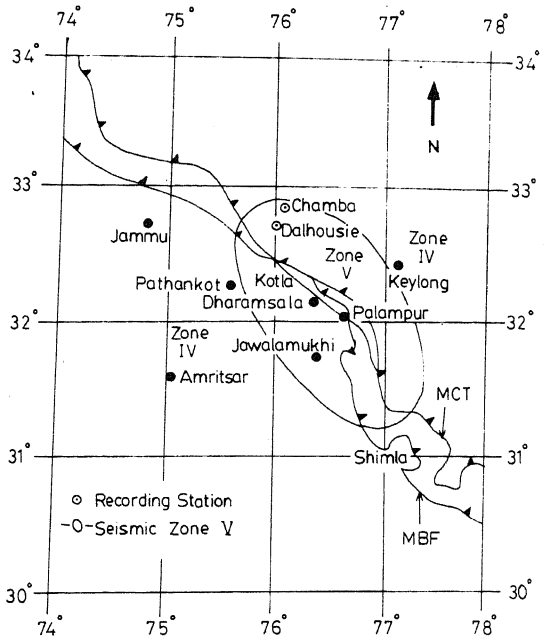


Figure 1. Map showing monitoring sites along with tectonic features in N–W Himalaya.

provides the facility to sum up any number of 15 min counting intervals for better counting statistics.

Radon monitoring in groundwater and soil-gas was carried out daily at Dalhousie station using the emanometry technique. The groundwater sample was collected from a 'bauli' (natural spring) at village Banikhet, 7 km from Dalhousie, in a sample bottle. A closed-circuit technique[11] was used to collect the gas that had passed through the water sample in the ZnS(Ag) detector cell as follows: The air was circulated in the closed circuit containing a hand-operated rubber pump, a water sample bottle, a drying chamber and a ZnS(Ag) detector cell for 10 min. Radon counts were recorded after 4 h, during which time the equilibrium between radon and its daughters gets established.

For measuring radon in soil-gas, auger-produced holes were left covered for 24 h, the same holes being used daily for radon monitoring. An emanometer was used to measure the alpha particle emission from radon in the gas fraction of a sample by pumping the soil-gas into a scintillation chamber using a closed-circuit technique. The alpha particles emitted from the decay of radon and its daughters impact the scintillator ZnS(Ag) creating an energy pulse in the form of photons. These photons were recorded by the scintillation assembly consisting of a photomultiplier tube and a scalar unit.

The Chamba and Dalhousie stations were set up as a part of radon network established in Himachal Pradesh to monitor the seismic activity of N–W Himalaya. Time series alpha-logger radon data for Dalhousie station are

plotted in Figure 2 for the period October 1993 to March 1995. The first radon anomaly was recorded on 26 October, with a radon value of 18,873 counts/day, 191% above the average value. This was followed by a seismic event of magnitude 1.3 M on 28 October within an epicentral distance of 21 km. Similarly, there were other radon anomalies followed by seismic events as shown by arrows in Figure 2 and listed in Table 1. This clearly shows that radon precursory signals are recorded prior to seismic events in N–W Himalaya.

Radon monitoring at Chamba started in July 1993. Time series radon data are plotted in Figure 3 with radon anomalies followed by seismic events (shown by arrows) as listed in Table 2. During March 1995, the radon emanation registered an average value of 4010 counts/day (Figure 4 *a*). On 13 March the radon peak occurred with radon high of 7600 counts/day. It started decreasing thereafter and fell below the average value after the Chamba earthquake of magnitude 5.1 M, which occurred on 24 March 1995 with its epicentre at Pliure, nearly 7 km away from Chamba town.

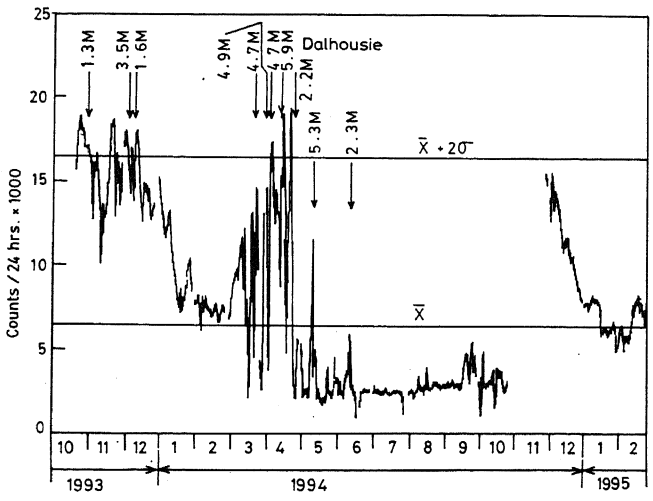


Figure 2. Time series alpha-logger data at the Dalhousie station.

Table 1. Correlation of radon anomalies at Dalhousie and the corresponding earthquake data

Date of radon anomaly	Date of the event	Magnitude	Alpha counts/day	Percentage increase above the mean value
26.10.93	28.10.93	1.3	18873	191
23.11.93	03.12.93	3.5	18625	187
02.12.93	05.12.93	1.6	18078	178
21.03.94	23.03.94	4.7	13166	103
24.03.94	30.03.94	4.9	14548	124
01.04.94	03.04.94	4.7	14575	124
06.04.94	09.04.94	5.9	17382	168
23.04,94	25.04,94	2.2	19343	198
10.05.94	10.05.94	5.3	11516	77
12.06.94	12.06.94	2.3	5846	–10

Radon monitoring at Dalhousie station shows some interesting trends (Figure 4 *b*): an enhancement till 16 March, then a fall, followed by a negative radon peak on 21 March, 3 days before the Chamba earthquake. Radon emanation showed a sudden rise after 21 March and this trend continued till 28 March; thereafter it started decreasing on 29 March. It reached its minimum value on 30 March after the seismic event of magnitude 3.0 M which occurred on 29 March. Chamba station became unoperational due to battery failure on 25 March and hence could not record the signal prior to 29 March event. The apparently different radon trends at both the stations may be explained on the basis of different geological conditions at the radon monitoring sites, their location from the main boundary fault (MBF) and the variation of the stress intensity factor[12].

Radon emanometry data (Figure 5) recorded the Chamba earthquake precursory signals more faithfully in both soil-gas and groundwater at Banikhet site near Dalhousie. The average radon values for March 1995 in soil-gas and groundwater were 71 and 98 counts/200 s, with a standard deviation of 97 and 103 counts, respectively. The radon anomaly was recorded simultaneously in both

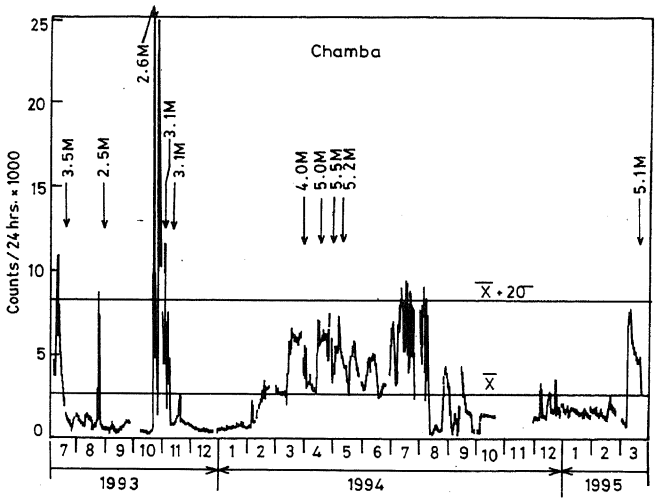


Figure 3. Time series alpha-logger radon data at the Chamba station.

Table 2. Correlation of the radon anomalies at Chamba and the corresponding earthquake data

Date of radon anomaly	Date of the event	Magnitude	Alpha counts/day	Percentage increase above the mean value
10.07.93	12.07.93	3.5	10906	309
26.08.93	28.08.93	2.5	8664	225
25.10.93	25.10.93	2.6	28837	982
04.11.93	10.11.93	3.1	11601	335
20.03.94	30.03.94	4.0	6611	148
18.04.94	18.04.94	5.0	7045	164
30.04.94	01.05.94	6.5	7536	182
09.05.94	09.05.94	5.2	7355	176

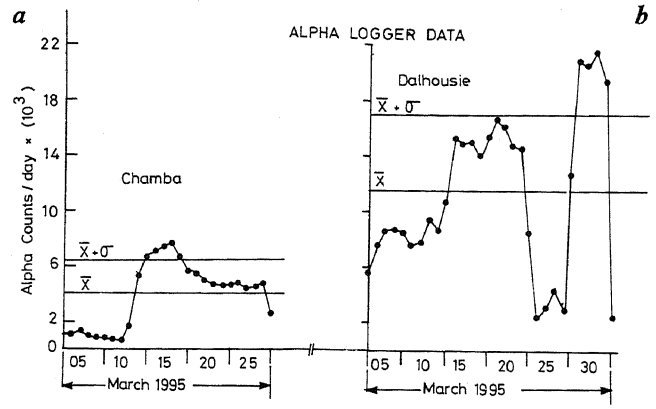


Figure 4. *a*, Alpha-logger radon data recorded at Chamba during March 1995. *b*, Alpha-logger radon data recorded at Dalhousie during March 1995.

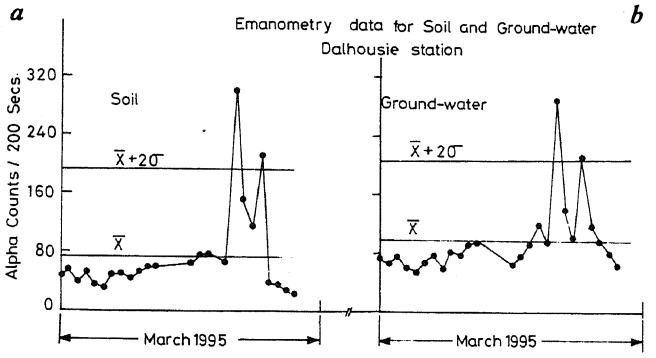


Figure 5. *a*, Radon emanometry data in soil-gas at Dalhousie during March 1995. *b*, Radon emanometry data in groundwater at Dalhousie during March 1995.

the media on 21 March, three days before the occurrence of Chamba earthquake and a day before a local event of magnitude 2.0 M, with peak values crossing the $x+2\sigma$ level. There was another peak on 24 March and then a sudden fall in the radon emanation rate after the strain was released. The simultaneous recording of radon peaks in both soil-gas and groundwater at the same site and under similar meteorological conditions before the occurrence of Chamba earthquake on 24 March establishes the efficacy of radon as an earthquake precursor. It also augurs well for setting up radon monitoring networks in N–W, Central and N–E Himalaya for the purpose of earthquake prediction studies and delineation of hidden faults.

1. Ulomov, V. I. and Mavashev, B. Z., *Dokl. Acad. Sci. USSR*, 1967, **76**, 319–322.
2. Birchard, G. F. and Libby, W. E., *J. Geophys. Res.*, 1979, **85**, 3100–310
3. King, C. Y., *J. Geophys. Res.*, 1980, **85**, 3065–3079.
4. Teng, T., Sun, L. and McRaney, J. L., *Geophys. Res. Lett.*, 1981, **8**, 441–444.
5. Wakita, H., *Science*, 1980, **207**, 882–883.
6. Hauksson, E. and Goddard, J. G., *J. Geophys. Res.*, 1981, **86**, 7037–7054.
7. Fleischer, R. L. and Mogro-Campero, A., *Geochim. Cosmochim. Acta*, 1985, **49**, 1061.
8. Virk, H. S. and Singh, B., *Nucl. Geophys.*, 1992, **6**, 287–291.
9. Virk, H. S. and Singh, B., *Tectonophysics*, 1993, **227**, 215–224.
10. Segovia, N., Meena, M. and Tamez, E., *Nucl. Tracks Radiat. Meas.*, 1993, **22**, 435–440.
11. Ghosh, P. C. and Bhalla, N. S., Proceedings of All India Symposium on Radioactivity and Meteorology of Radionuclides, A.E.E.T., Bombay, 1966, pp. 226–239.
12. Hauksson, E., *J. Geophys. Res.*, 1981, **86**, 9397–9410.

ACKNOWLEDGEMENTS. We acknowledge the financial support of the Department of Science and Technology (DST), New Delhi, under Himalayan Seismicity Project No. DST/23 (33)/ESS/91. We are also grateful to the Director IMD, New Delhi, for the supply of seismic data and Dr V. C. Thakur, Director, Wadia Institute of Himalayan Geology, Dehradun, for helpful discussions.

Received 13 June 1995; revised accepted 11 July 1995

Palaeogeographic evolution of the Akkulam lake and its geotechnical implications in the urban development of Thiruvananthapuram

K. Reghunadh, K. Sajan and S. Mohanakumar*

Marine Geology Division, School of Marine Sciences, Cochin University of Science & Technology, Cochin 682 016, India
*Govt. Engineering College, Thiruvananthapuram 695 081, India

The palaeogeographic evolution of the Akkulam lake and its geotechnical significance based on field observations, mapping, lithological core sampling and geotechnical laboratory studies are highlighted. The findings have important implications in the future urban development of Thiruvananthapuram.

FROM a geological perspective all lakes are temporary features in the terrestrial landscape. Lakes originate, evolve and ultimately disappear. They differ from one another mostly in terms of their mode of formation, course of evolution, duration of existence and their ultimate disappearance.

The present investigation is concerned with the palaeogeographic evolution of the Akkulam lake and its geotechnical implications in the future development of the fast-growing urban area of the city of Thiruvananthapuram.

The Akkulam lake is one of the minor lakes located within the coastal plain immediately north of Thiruvananthapuram. The lake presently covers an area of 1.7 km^2 and is intermittently connected with the Lakshadweep

sea through a tidal channel. A temporary transverse sand bar, generally 3 m wide and over 15 m long, cuts off the tidal channel from the sea, preventing direct contact with the latter during most of the year, except during periods of high rainfall. Thus, the lake forms a 'blind' estuary[1]. The manmade inland canal, the Parvathy Puthen Ar, enters this tidal channel. This canal continues further southwest into the coastal plain at a distance of 1.5 km from the coastline, almost parallel to the strandline. The two bridges, one connecting the coastal roads and the other connecting the railway tracks, are constructed across this tidal channel (Figure 1).

The study comprised the following methodology.

1. Geological and geomorphological mapping of the region in and around the Akkulam lake covering an area of 20 km^2 in the scale 1 : 20,000.
2. Drill core sampling, field study of sites of excavations, burrow pits, etc.
3. Lithological study of drill cores, materials of excavations (well cuttings, burrow pits and pits for foundations, etc.).
4. Geotechnical study of the lithological samples.
5. Preparation of a map showing zones of geotechnical hazards.

The field studies conducted in the region surrounding the basin of the lake have indicated that the lake was more extensive in the past, and an arm of the lake appears to have extended up to Kannanmoola for a distance of 6 km from its present margin (Figure 2).

The geomorphological features of the southeastern part of the study area adjoining the lake afford clear evidence of the shrinking of the lake basin to its present size. The landform is characterized by a flat, featureless terrain not exceeding 2.5 m above the water surface of the lake. This low-relief landform resembles that typical of flood plains of aggradation or of valley fills. Data from drill holes, pits and excavations collected from nine selected sites within the valley floors between Akkulam and Kannanmoola afford unmistakable evidence to show that it is genetically a valley flat of aggradational nature where presence of abnormal thickness of sediments has been shown to exist. The sediments forming the infilled portion within the area between the hill-side escarpments and slopes have an aggregate thickness ranging from 11 to 20 m. The width of this zone of aggradation is not uniform, with widths varying from 100 to 500 m. The total area occupied by the sediment infilled zone is estimated to be 11,80,000 m^2. Considerable thickness (ranging in aggregate thickness from 6 to 22 m) of plastic clays alternating with sediments of silty and sandy horizons occupies this tract (Table 1).

A narrow channel of 15 m width occupies the central portion of the sediment-infilled portion of the lake. While the net flow of water in this channel is seaward, the direction of flow is reversed once or twice daily according to the tidal regime. Tidal influence is experienced in this channel up to just north of Vempalavattom and perceptible effects are recorded during some months in this channel as far as Kannanmoola, when the sand bar of the estuarine mouth is breached. Marshy tracts and paddy fields occupy both sides of the sediment-filled portion of the Akkulam basin, extending laterally up to the terraced bluffs flanking the valley-side escarpments, where extensive coconut cultivation prevails. In the neighbourhood of Kannanmoola the sediment infilled part of the Akkulam basin supports extensive marshy tracts almost throughout the year, where enormous thickness of plastic clays highly enriched in organic matter is known to attain an aggregate thickness of over 20 m.

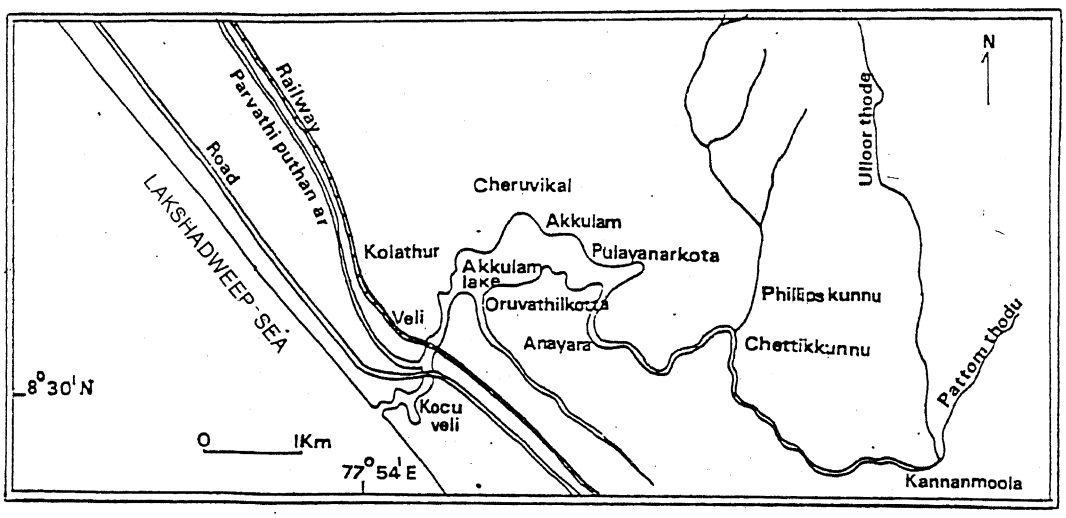


Figure 1. Geological map of the region around the Akkulam lake covering an area of 20 km^2.

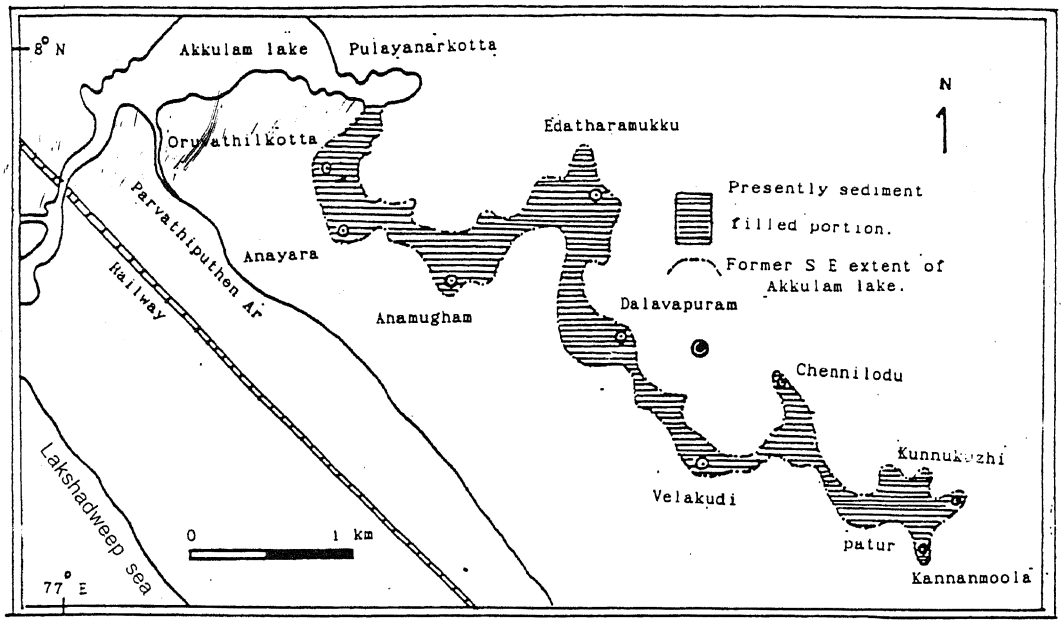


Figure 2. Geomorphological map of the region around the Akkulam lake covering an area of 20 km^2.

This locality is situated 5 km inland from the present coastline and occupies almost the central part of the city of Thiruvananthapuram.

The drill core samples collected from the sediment-filled portions of the Akkulam basin from Kannanmoola comprise occasional horizons containing skeletal fragments and complete shells of molluscs, and other invertebrates characteristic of brackish water environments. Thus, the palaeontological observations afford another strong evidence for the prevalence of brackish water condition up to Kannanmoola somewhere in the recent past.

Yet another evidence supporting the palaeogeographic change of the Akkulam lake basin is afforded by the drainage pattern of the existing fluvial system physically associated with the present lake basin. A number of stream systems extending over an aggregate area of about 30 km debouch their discharges into the Akkulam lake. Among these, the southerly flowing Ulloor Thode and southwesterly flowing Pattom Thode are the prominent ones. In addition, a few smaller streamlets enter the lake from its northern side traversing Kolathur, Cheruvickal and neighbouring localities. The Ulloor Thode and the Pattom Thode combine together at the centre of the marshy tract of Kannanmoola – the locality, as already noted, represents probably the former southeastern extremity of the present Akkulam lake. The sediment accumulation process in the portion of the lake basin from Kannanmoola up to Akkulam left only a narrow channel up to its present margin to drain away the discharges received from Ulloor Thode and Pattom Thode. A close study of the drainage pattern of the region reveals an abrupt northwestward diversion of the downstream portion of the stream channel beyond the point of confluence of Pattom Thode and Ulloor Thode. This unusual stream pattern indicates perhaps that in the past the streams had their termination near Kannanmoola up to where the southeast-trending arm of the Akkulam lake extended and subsequent infilling of this portion of the lake basin forced the stream channel to extend gradually up to Akkulam – the lake's present margin.

The urban area of Thiruvananthapuram is growing at an accelerated rate and the urban–suburban boundary is expanding at different rates along different directions. The scarcity of land for building construction has forced increasing reclamation of most of the paddy fields that occupied a substantial portion of the urban area. The building construction activity has already reached the infilled southeastern portion of the Akkulam basin. This tract being a geotechnically hazardous zone, special precautionary measures should be taken for foundation treatment while planning construction of buildings, especially high-rise constructions. Field studies have indicated serious damages of various sorts even in newly constructed buildings in this tract due to foundation problems. The enormous thickness of plastic clays in the sedimentary succession of this tract, extending up to Kannanmoola should be taken into consideration in planning future urban development of the infilled portion of the Akkulam basin.

The clay-enriched sedimentary succession has a significant influence upon its geotechnical properties. The moisture content and the energy with which this moisture

Table 1. A representative lithological borehole log at Velakudi

	Lithology	Gravel %	Sand %	Silt %	Clay %
A	Clayey and sandy silt	–	12	42	46
B	Stiff lateritic clay	8	30	52	10
C	Soft grey clay with peaty matter	–	4	6	90
D	Carbonaceous plastic clay	–	–	8	92
E	Sandy and silty clay	–	12	18	70
F	Carbonaceous plastic clay with occasional peaty horizons	–	3	11	86
G	Sandy clay	–	18	20	62
H	Silty clay with shell fragments	–	2	30	68
I	Medium sand with shell fragments	–	42	26	32
J	Disintegrated rock	16	52	32	–
K	Hard rock (Khondalite)				

is held are the important factors to be taken into account while planning high-rise buildings in such areas, since these determine the consistency, strength and volume changes resulting from swelling, shrinkage and consolidation, which are also related to permeability and migration of moisture. Volume changes can occur as a result of loading consequent to consolidation and heave, due to reduction in void ratio. Because of their low permeability, the rate of consolidation will be a relatively slow process but the net effect will be considerable over a long span of time. The sensitivity of the clays of the region, i.e. the ratio of the undisturbed to the remoulded strength at the same moisture content ranges from about 3 to 6. Because of the high content of moisture of the clays, any fluctuation in the moisture content may increase the sensitivity of these clays. However, the ultimate bearing capacity of foundations on clay-enriched successions is influenced by its shear strength, which, in turn, depends on its consistency as well as the shape and depth at which the structure is placed. There are unmistakable evidences of severe foundation problems in the sediment-infilled portion of the Akkulam basin in the form of collapse of bridges due to tilting of foundations at localities such as Anamughom, near Edatharamukku and Nelliazhi. The shear strength of the clays determined for samples obtained from deep excavations and drill cores gave values ranging from 20–60 kN/m^2.

The study revealed that the Akkulam lake initially extended from its present margin up to Kannanmoola in southeast. Geomorphological, geological and palaeontological evidences support this view. This infilled arm of the lake contains considerable thickness of plastic clays and thus constitutes a geotechnically hazardous zone. This fact should be taken into consideration while planning major civil engineering constructions for the future urban development programmes.

1. Day, J. H., *Trans. R. Soc. S. Afr.*, 1951, **33**, 53–91.
2. Bell, F. G., *Engineering Geology and Geotechnics*, Newnes–Butterworths, London, 1980.

Received 26 December 1994; revised accepted 25 July 1995

Characterizing intergeneric regenerants of protoplast fusion between *Hyoscyamus muticus* (Egyptian henbane) and *Atropa belladonna* (Indian sagangur)

C. C. Giri and P. S. Ahuja*

Centre for Plant Molecular Biology, Department of Genetics, Osmania University, Hyderabad 500 007, India
*Plant Tissue Culture Division, Central Institute of Medicinal and Aromatic Plants, Lucknow 226 015, India

In the absence of a selection system, protoplast fusion between *Hyoscyamus muticus* L. and *Atropa belladonna* L. was obtained. The conditions for obtaining such a fusion were determined. The regenerants obtained subsequent to fusion were characterized based on morphology and cytological analysis. The regenerants exhibited phenotypes that were intermediate or resembling the parental types.

HYOSCYAMUS MUTICUS L. was introduced into India from Egypt for its cultivation and medicinal use[1]. Combining some characteristics of this species, viz. alkaloid content and arid habitat in a plant like *Atropa belladonna* (a sexually incompatible distant relative) is an essential step in the improvement of these medicinally important plants. Production of somatic hybrids was reported between sexually incompatible plant species specifically for obtaining higher content of secondary metabolites[2]. Besides somatic hybridization the improvement of plant bearing tropane group of alkaloids was reported through protoclonal variation[3]. Protoplast fusion between *A. belladonna* and *H. muticus* was attempted and the analysis of a few of the fusion products is reported in the present paper.

Seed materials of *A. belladonna* L. ($2n = 72$) and *H. muticus* L. ($2n = 28$) were obtained from CIMAP experimental farm. Callus cultures were established from leaf explants on MS[4] medium supplemented with α-naphthaleneacetic acid (NAA, 2 mg/l) and benzylaminopurine (BAP, 0.5 mg/l) from *in vitro* grown plants. Suspension cultures from both the plants were established using the same medium in liquid[5]. Protoplasts were isolated from 3-month-old suspension cultures. Parental protoplasts (1×10^6 protoplasts/ml of each parent) were mixed in a centrifuge tube, centrifuged for 5 min at $30 \times g$ and the supernatant removed. One ml of fusion solution containing 25% PEG (mol. wt. 3500), 9% mannitol, 2.36% $Ca(NO_3)_2 \cdot 4H_2O$ was added slowly along the wall of the centrifuge tube and then incubated for 0–10 min at 30–32°C. PEG was diluted out slowly with a high-pH Ca^{2+} solution (0.74% $CaCl_2 \cdot 2H_2O$, 0.375% glycine, 10% mannitol, pH 10.5). The duration of incubation in the fusogen or dilutant was varied to determine the optimum fusion frequency.

The protoplasts were then further diluted with a stabilizer solution (0.74% $CaCl_2 \cdot 2H_2O$, 0.75% glycine, 10% mannitol in CPW salts[6], pH 5.8) for at least 1 h. Finally, the settled protoplasts were washed with protoplast culture medium (PCM). Estimation of fusion frequency was done as per Kinsara *et al.*[7] using fluorescence labels FITC (fluorescein isothiocyanate), RITC (rhodamine isothiocyanate)[6], before culture of protoplasts.

After fusion, the heterokaryons along with the parental protoplasts were cultured at an optimum culture density (1×10^5 protoplasts/ml) in a standardized PCM (MS + 3% sucrose (w/v), 9% mannitol (w/v), 1 mg/l NAA, 0.2 mg/l 2,4-dichlorophenoxyacetic acid (2,4-D), 0.5 mg/l BAP)[5].

Chromosomal counts in the actively growing healthy root tips (1–2 mm) of plants regenerated *in vitro*, hybrid callus and parental seedlings were determined. The roots or calli were pretreated in a saturated solution of *p*-dichlorobenzene (PDB) for 4 h (initially for 5 min in a deep freezer and subsequently at 7°C). The tissue of *A. belladonna* was pretreated in water for 24 h at 2°C. All the tissues were fixed in 1 : 3 :: acetic acid : absolute alcohol. The samples were stained with a mixture (1 : 9) of 1 N HCl and 2% aceto-orceine (dissolved in 45% glacial acetic acid).

Protoplast fusion between *H. muticus* and *A. belladonna* was confirmed on the basis of differential staining by FITC and RITC. The frequency of fusion of protoplasts from *H. muticus* and *A. belladonna* is shown in Table 1. The frequency of the heterokaryon formation with PEG alone is low (0–0.9%). Heterokaryon formation was 2.6% when the parental protoplasts were incubated in high Ca^{2+}/high pH solution for 20 min. However, this frequency could be marginally increased to 3.0% if the parental protoplasts were incubated in PEG for 5 min earlier. Also bursting of protoplasts was noticed to be lower in this treatment than with PEG alone. Thereafter,

Table 1. The effect of the interaction of PEG and ionic solution (high pH/Ca^{2+}) on heterokaryon formation* between protoplasts of *A. belladonna* labelled with FITC and *H. muticus* labelled with RITC

Incubation time (min) in PEG	Incubation time (min) in high pH/Ca^{2+}					
	0	15	20	25	30	45
0	0	0.05	2.6	0.96	0.06	— (B)
2	0.08	0.32	0.9	0.12	0.001	— (B)
4	0.06	0.25	1.86	0.08	0.015	— (B)
5	0.01	0.52	3.0	0.12	0.001	— (B)
8	0.09	0.1	0.08	0.03	0 (B)	— (B)
10	0.02	0.09	0.07	0.06	0 (B)	—

*Fusion frequency % (Percentage of heterokaryons per total number of protoplasts).
B = Bursting of protoplast.
Observations were taken after 1 h of fusion process was over under UV using a fluorescence inverted microscope.
A minimum of 25 microscopic fields were observed to arrive at the fusion frequency.

incubating the protoplasts for further time periods in either of the solutions did not increase the fusion frequency. Plating efficiency of parental protoplasts was to the tune of $56 \pm 2\%$ and $47 \pm 1.5\%$ in case of *H. muticus* and *A. belladonna*, respectively. However, when protoplasts were cultured after fusion at same culture density as parents (1×10^5 protoplasts/ml), the plating efficiency was observed to be $68 \pm 2\%$. Fast-growing colonies were visible within 4–5 weeks of culture. From 50 fast-growing colonies 15 callus lines were obtained. These callus lines were then transferred to a semisolid agar (0.8%) MS regeneration medium containing 3% sucrose 0.1 mg/l NAA and 0.5 mg/l BAP. Regenerated shoots developed into complete plantlets when the shoots were transferred to rooting medium containing half-strength MS nutrients without phytohormones. The regenerated plants were selected visually on the basis of the morphology, pattern and plant regeneration frequency. Calli obtained after fusion were highly compact and green compared to parental calli, which were fragile and yellowish green in nature. There was significant difference in the regeneration potential of hybrid callus compared to parental protoplast callus. The regeneration frequency from protoplast-derived calli of *H. muticus* and *A. belladonna* protoplasts was 80% and 60%, respectively. However, the regeneration frequency of callus obtained after fusion was extremely high (100%), i.e. whole of the callus would transform into regenerants. Similar screening for somatic fusion product was reported in other plant species where no selection system was available[8–11].

Amongst the different callus lines studied, plants regenerated from four selected calli were peculiar as some regenerants showed intermediate phenotype (Figure 1). From four such callus lines, 15 plants showing intermediate phenotypes were obtained and 9 of these plants were transferred to glasshouse. These plants flowered within 9–10 weeks of transfer to soil. The flower morphology in such plants was different from the parental types. The flowers had larger petal

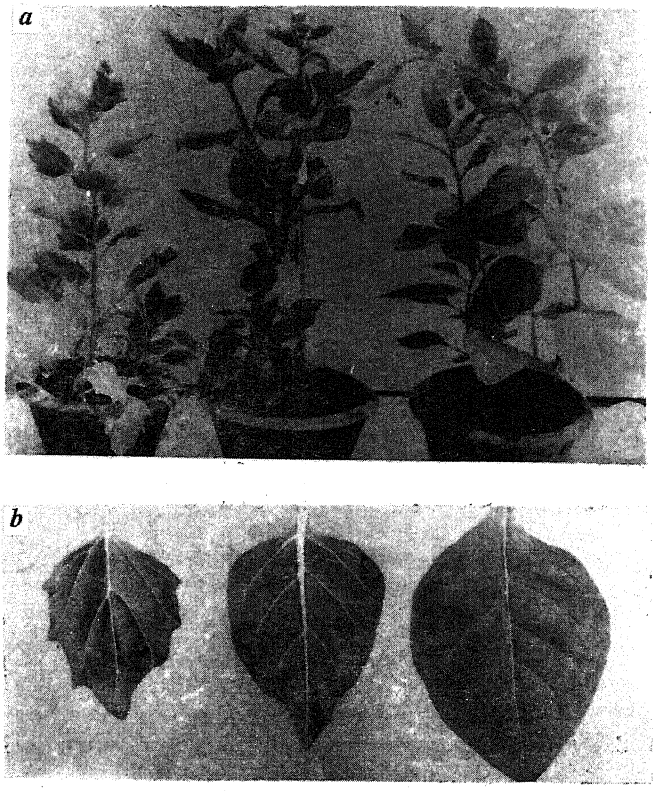


Figure 1. Morphological variation amongst fusion-derived regenerants as compared to parents. *a*, *H. muticus* (left), *A. belladonna* (right) and fusion-derived regenerant (middle); *b*, Leaf morphology of *H. muticus* (left), *A. belladonna* (right) and fusion-derived regenerant (middle).

Table 2. Chromosome analysis of the calli obtained after fusion between *A. belladonna* ($2n = 72$) and *H. muticus* ($2n = 28$) and their regenerants

Replicate no.	Callus lines				Roots* of regenerated plants		
	1	2	3	4	Symmetric phenotype	Asymmetric phenotype	
						A. belladonna	*H. muticus*
1	110	65	105	92	100	81	40
2	105	87	140	95	105	86	56
3	102	86	130	101	102	70	30
4	107	83	125	107	103	56	25
5	85	82	97	100	100	96	28
6	—	120	100	—	103	100	42
7	—	100	—	—	100	53	24
8					108	62	28
9					104	72	32
10					—	73	38
11					—	75	28
12					—	82	39
13					—	72	28
14					—	—	28

*A minimum of 10 samples for each tissue were investigated for chromosome analysis.

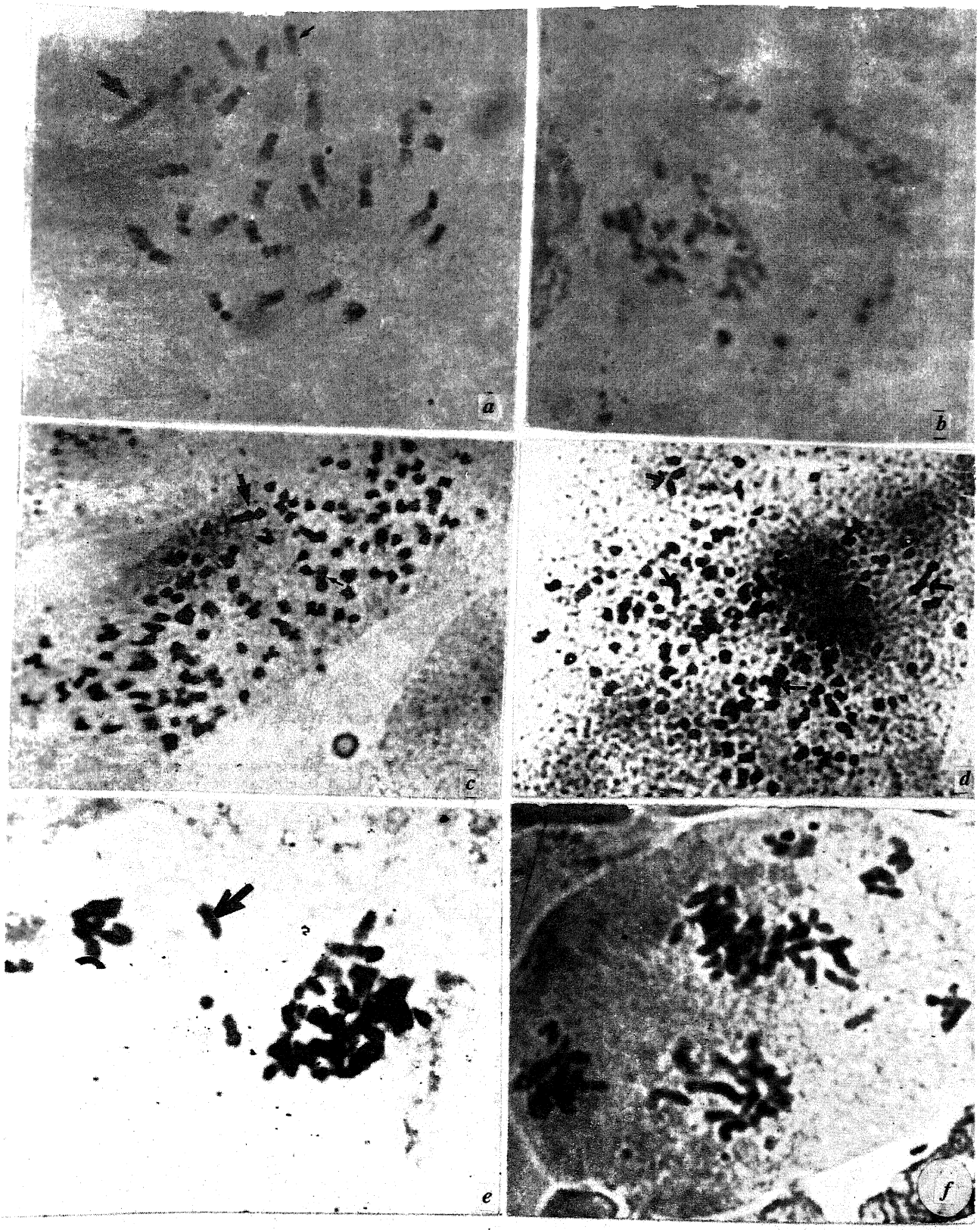


Figure 2. Somatic chromosome analysis of the parents and putative asymmetric hybrids between *H. muticus* and *A. belladonna*. *a,b*, Metaphase chromosome of *H. muticus* ($2n = 28$) and *A. belladonna* ($2n = 72$). *c,d*, Chromosomes of fusion-derived regenerated plants ($2n = 100, 109$), majority of chromosomes are smaller *belladonna* type with few *muticus* type (arrow). *e*, Unequal anaphase separation (disjunction) with laggard chromosome (arrow). *f*, Multipolarity of chromosome during anaphase separation.

size compared to parents. They also differ in number of petals (6) and stamens (6) compared to parents (5). The shape and colour of anther was changed to heart-shaped white compared to rod-shaped purple in the parents.

These four selected callus lines in subsequent culture passage gave rise to two types of regenerants resembling *H. muticus* and *A. belladonna*. The regenerants included 12 asymmetric hybrids resembling *A. belladonna* and 17 resembling *H. muticus*. In a similar finding both *Arabiodopsis* and *Brassica* plants were found to regenerate from the same hybrid callus during the culture passage[12].

Table 2 shows mitotic chromosome analysis of selected calli and regenerated shoots showing intermediate phenotype (symmetric hybrids) or resembling the parental types (asymmetric hybrids). In the hybrid callus the chromosome number ranged from 83 to 140. Few regenerants which possessed intermediate phenotype showed chromosome numbers between 100 and 108. Of the asymmetric hybrid plants, those resembling *A. belladonna* possessed 56 to 100 chromosomes and those resembling *H. muticus* was 24–56. The cytological analysis of the chromosomes for the plants resembling *A. belladonna* showed a few larger-sized chromosomes of *H. muticus* (Figure 2 *c*, *d*). The *H. muticus* type of regenerants showed the presence of smaller-sized chromosomes of *A. belladonna*. This could have resulted in the production of asymmetric hybrids. Cytological anomalies such as elimination of chromosomes, unequal separation of chromosomes with the presence of laggards and multipolar orientation of chromosomes during anaphase were observed when the calli were observed after several subcultures (Figure 2 *e*, *f*).

The chromosome elimination subsequent to somatic hybridization was found to occur in most of the fusion studies. Chromosome elimination was observed in fusion involving *Arabiodopsis thaliana* and *Brassica campestris*[13]. In a study on the fusion between *A. belladonna* and *N. chinensis* the elimination of parental chromosomes has been observed. In most of the cases species-specific elimination takes place[14–16]. Similar findings were also obtained in studies on somatic hybridization between *A. belladonna* and *Datura innoxia*[17]. The phenomenon of chromosome elimination in the present study can be exploited for raising asymmetric hybrids and may be used for limited gene transfer as has been reported earlier with other crops[18]. The present work indicates the possibility of producing an intergeneric somatic hybrid between two medicinally important plants in the absence of a selection system.

1. Hussain, A., Singh, P. and Singh, A., *Indian J. Pharm.*, 1979, **41**, 46–48.
2. Schieder, O., *Mol. Gen. Genet.*, 1978, **162**, 113–119.
3. Giri, C. C. and Ahuja, P. S., *Curr. Sci.*, 1994, **66**, 445–448.
4. Murasihige, T. and Skoog, F., *Physiol. Plant.*, 1962, **15**, 473–497.
5. Giri, C. C. and Ahuja, P. S., *Indian J. Exp. Biol.*, 1990, **28**, 249–251.
6. Frearson, E. M., Power, J. B. and Cocking, E. C., *Develop. Biol.*, 1973, **33**, 130–137.
7. Kinsara, A., Patnaik, S. N., Cocking, E. C. and Power, J. B., *J. Plant Physiol.*, 1986, **125**, 225–234.
8. Belliard, G., Pelletier, G. and Ferault, T. M., in Proceedings of the 8th Congress, Madrid, 1977, pp. 237–242.
9. Scowcroft, W. R. and Larkin, P. J., *Theor. Appl. Genet.*, 1981, **60**, 179–184.
10. Rosen, B., Hallden, C. and Heneen, W. K., *Theor. Appl. Genet.*, 1988, **76**, 197–203.
11. Sikdar, S. R., Chatterjee, G., Das, S. and Sen, S. K., *Theor. Appl. Genet.*, 1990, **79**, 561–566.
12. Hoffmann, F. and Adachi, T., *Planta*, 1981, **153**, 586–593.
13. Gleba, Y. Y. and Hoffmann, F., *Planta*, 1980, **149**, 112–117.
14. Kao, K. N., *Mol. Gen. Genet.*, 1977, **150**, 225–230.
15. Tobacizadeh, Z., Prennes, C. and Bergounioux, C., *Plant Cell Rep.*, 1985, **4**, 7–11.
16. Endo, T., Kamiya, T., Masumitsu, Y., Morikawa, H. and Ymada, Y., *J. Plant. Physiol.*, 1988, **129**, 453–459.
17. Krumbiegel, G. and Schieder, O., *Planta*, 1981, **153**, 466–470.
18. Glimelius, K., *Plant Cell Tiss. Organ Cult.*, 1988, **12**, 163–172.

ACKNOWLEDGEMENTS. The work was supported by Council of Scientific and Industrial Research (CSIR), New Delhi. Laboratory facility by Director, CIMAP, Lucknow is acknowledged.

Received 30 May 1994; revised accepted 18 July 1995

Mechanism of wound healing induced by chitosan in streptozotocin diabetic rats

A. Naseema, Pius S. Padayatti and C. S. Paulose

Molecular Neurobiology and Cell Biology Unit, Centre for Biotechnology, Cochin University of Science & Technology, Cochin 682 022, India

Diabetes mellitus is a common metabolic disorder associated with many pathological conditions. Prolonged or incomplete wound healing is one among them. This is due to poor blood circulation, alterations in circulating constituents, abnormal platelet aggregation and decreased fibrinolytic activity. The aim of the present study was to envisage the role of chitosan in wound healing and to investigate its effect in normalizing the aggregation pattern of platelets in diabetes. At very low concentrations, chitosan has no effect on platelet aggregation, while high concentrations can aggregate the platelets. Chitosan is not capable of reversing the increased rate of aggregation in diabetic platelets. The number of platelets which may affect the normal clotting process were also significantly lower in diabetic rats. Chitosan was found to be a good wound-healing agent, even though it does not reverse the altered platelet aggregation. Chitosan may be enhancing wound healing by some other mechanism independent of its effect on platelet aggregation.

DIABETES mellitus is caused by decreased secretion of insulin by the β cells of Islets of Langerhans, lack of

glucose-metabolizing enzymes and gluconeogenesis from noncarbohydrate sources. This disease is associated with many pathological conditions. Prolonged or incomplete wound healing is one among them. This is due to poor blood circulation, alterations in circulating constituents, abnormal platelet aggregation and decreased fibrinolytic activity. Platelet aggregation is of primary importance in normal haemostasis. When a blood vessel is cut, a platelet plug is formed by the process of platelet aggregation[1]. So any change in platelet function or in platelet number can finally lead to a difference in the blood-clotting process, which, in turn, causes prolonged wound healing[2]. Poor blood circulation makes diabetics prone to infection of the foot, commonly called diabetic gangrene.

In some recent studies, agonists like ADP, hydrogen peroxide, thrombin, etc., were used to understand the process of platelet aggregation. Many workers reported the effect of chitosan as an enhancer in wound healing. Chitosan in solution coagulates red cell membranes. It is this property of chitosan which finds wide applications in the field of neurosurgery and wound healing[3].

In the present study our aim is to envisage the role of chitosan in wound healing in diabetes and to investigate the effect of chitosan in normalizing the altered platelet aggregation in diabetes mellitus.

Animals used were Sprague Dawley strains of rats with ~ 200 gm body weight. They were maintained on gram (*Cicer arietinum*) and tap water supplied *ad libitum*.

For wound-healing studies, female rats were used. They were divided into group I (control rats), group II (control rats with chitosan powder applied to artificial wounds), group III (diabetic rats), group IV (diabetic rats with chitosan powder applied to artificial wounds).

Animals used for platelet aggregation studies were classified into group I (control animals) and group II (diabetic animals).

Diabetes was induced using streptozotocin (STZ)[4]. Body weight and blood glucose levels were measured throughout the experimental period. Glucose level from tail vein blood was estimated with GOD-PAP method by using a glucose kit (Merck).

In order to examine the process of wound healing, two artificial wounds, one on the back of the neck and the other on the leg, were made in the above four groups of animals by removing the cutis and subcutis (2–3 cm long). Then these artificial wounds in group II and group IV animals were treated with chitosan powder and the other two groups were left as such. Wound healing was assessed visually.

Platelets were prepared by collecting blood from tail vein of control and diabetic rats using the method of Domino *et al.*[5]. Platelets were counted using haemocytometer under phase contrast microscope[6]. Platelet aggregation was assayed using the method of Jamaluddin and Krishnan[7]. The PRP was diluted to make it 7×10^7 cells/ml of the experimental solution. The agonist used was 1 mg/ml chitosan solution. The agonist concentration used varied from 15 to 60 µg.

For finding out the initial rate of platelet aggregation, the number of single platelets was counted by withdrawing the sample at zero time and at various intervals after adding the agonist from a cuvette kept outside the spectrophotometer and counted immediately for finding out the percentage of single platelets[6].

Streptozotocin administration caused an elevation in the blood glucose level, decrease in body weight and increase in water uptake in all experimental animals (group II) (Table 1).

At very low concentrations of the agonist (3–15 µg/ml), the turbidity was found to increase as a function of time up to a particular point, which then stabilized in both control and diabetic samples (Figure 1).

At high concentrations of chitosan (50–60 µg/ml), the turbidity increased first and then decreased rapidly and stabilized. The normal agonist epinephrine also showed the same pattern of aggregation (data not shown). In diabetic samples the aggregation started within 10 s,

Table 1. General characteristics of experimental rats

	Blood glucose level (mg/dl)	Body weight (g)	Water uptake (ml)
Control	96.4 ± 3.0	170.0 ± 3.5	17.7 ± 1.1
Diabetic	256.9 ± 26.0*	147.5 ± 7.3*	61.0 ± 6.7*

Values are mean ± SEM of 5–6 separate experiments.
*$P < 0.05$ compared to control.

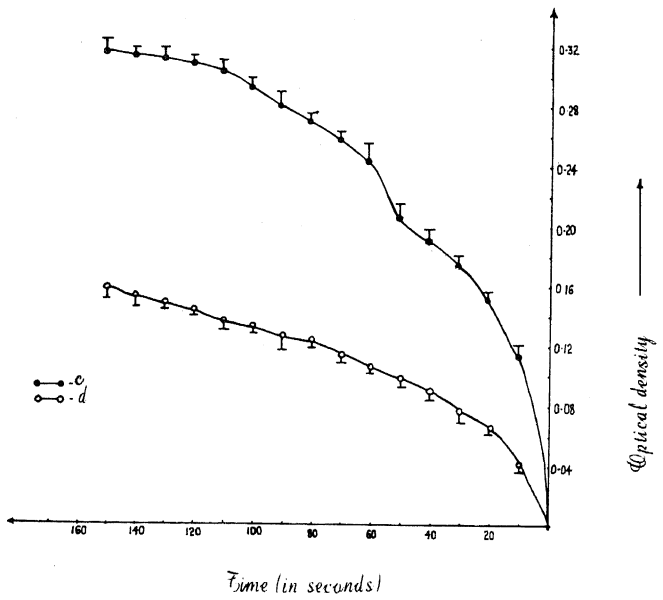


Figure 1. Platelet aggregation pattern induced by low concentrations of chitosan in control (●) and diabetic (O) rat platelet-rich plasma. Values are mean ± SEM of 4–6 separate experiments.

while in control it started within 20 s (Figure 2). The decrease in turbidity was associated with a decrease in the percentage of single platelets. The initial rate of aggregation (r_0) obtained showed significant increase ($P < 0.05$) in diabetic samples compared to controls (Figure 3).

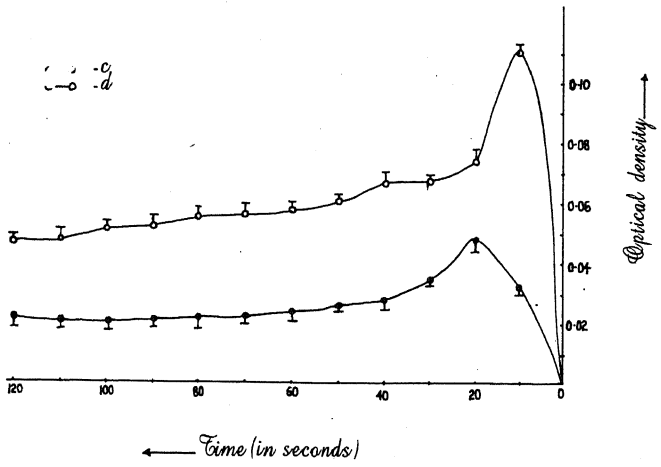


Figure 2. Platelet aggregation pattern induced by high concentrations of chitosan in control (●) and diabetic (O) rat platelet-rich plasma. Values are mean ± SEM of 4–6 separate experiments.

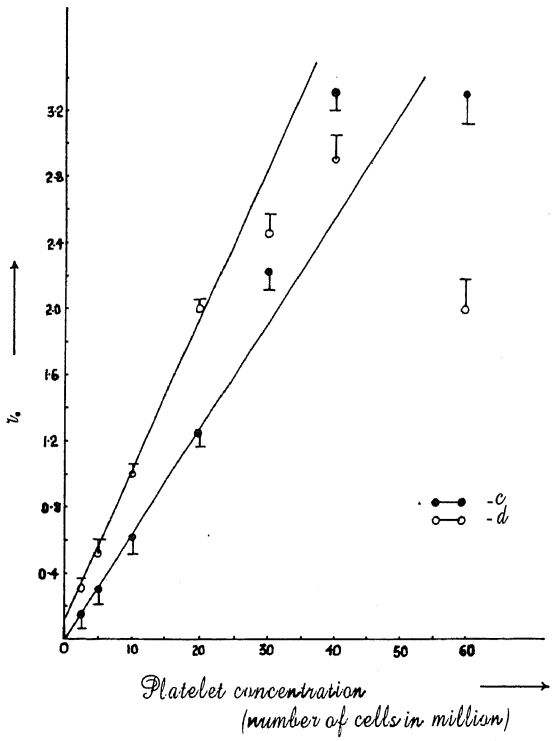


Figure 3. Changes in spectrophotometrically determined initial rates (r_0) of chitosan-induced aggregation of control (●) and diabetic (O) rat platelet-rich plasma as function of platelet concentration. The lines were drawn to show deviation of data from linearity. r_0 was calculated from the graph of percentage of single platelet (on Y axis) against time (on X axis). The slope of the graph is expressed as r_0.

The number of platelets in diabetic rats was significantly less ($P < 0.05$) compared to control rats (Table 2).

Visual observation of chitosan applied on artificial wounds in experimental animals showed considerable difference in wound-healing capability compared to controls without chitosan treatment. The chitosan-applied wounds in diabetic and control rats healed much earlier than the wounds without chitosan treatment (Table 3).

The significant elevation of blood glucose level, decrease in body weight and increase in water uptake in STZ-induced diabetes mellitus is due to altered carbohydrate metabolism brought about by selective destruction of Islet β cells by the toxin[8].

From our findings it was observed that very low concentrations of chitosan did not enhance platelet aggregation. From their study Olsen *et al.*[3] reported the mechanism by which chitosan induces gel formation with RBC, shown to be due to interaction of the positively charged chitosan polymers with receptors containing neuraminic acid residues on the cell membrane. The mechanism behind the interaction between platelets and chitosan may be in the same pattern. At very low concentrations there may be inadequate binding sites for chitosan. This may be the reason why low concentrations of chitosan did not enhance platelet aggregation.

We observed enhanced platelet aggregation with high concentrations of chitosan. The optical density, which

Table 2. Platelet counts in diabetic and control rats

	Animal status (Numbers/ml of blood × 10^8)	
	Male	Female
Control	2.45 ± 0.16	3.16 ± 0.26
Diabetic	1.69 ± 0.04*	1.87 ± 0.10*

Values are mean ± SEM of 4–5 separate experiments.
*$P < 0.05$ compared to control.

Table 3. Visual observation of wound healing in control and experimental rats

Days	Control	Diabetic	Control + chitosan	Diabetic + chitosan
1	–	–	–	–
2	–	–	–	–
3	–	–	–	–
4	–	–	–	–
5	–	–	–	–
6	–	–	+	–
7	–	–	+	+
8	–	–	+	+
9	–	–	+	+
10	–	–	+	+
11	–	–	+	+
12	–	–	+	+
13	+	–	+	+
14	+	–	+	+

–, No healing.
+, Complete healing.

is a direct measure of turbidity, increased first and then decreased. Many workers have reported the pattern of ADP-induced platelet aggregation[9]. The pattern of chitosan-induced aggregation resembles the ADP-induced pattern. Thus, chitosan can act as a normal agonist like ADP. The initial increase in optical density was reported to be due to shape change of platelets[10].

In the present study we observed increased platelet aggregation in diabetic condition. This can be explained by the findings of many workers. In 1972, Kwaan *et al.*[11] found out that there is significant increase in ADP-induced platelet aggregation in diabetic samples compared to controls. Aggregation in washed platelets from rats with diabetes was enhanced; this results in prolongation of thrombogenesis[12]. Platelet aggregation in response to ADP in washed platelets was increased in diabetic animals compared to controls. This is due to increased platelet size in diabetic animals[13]. *In vitro* platelet hyperaggregation was found induced with epinephrine in diabetic condition[14].

The observed significant decrease in the number of platelets in diabetic rats can be explained as follows. Zucker explains in his study that in severe atherosclerosis and in diseases such as thrombosis and diabetes, in addition to the growth factor, two other substances, platelet factor-4 and β thromboglobulin are released in large amounts, resulting in their rapid depletion from the circulation[15]. So the observed decrease in platelet number in both sexes of animals during diabetes may be related to altered secretory function of platelets. The decreased platelet number directly affects the platelet plug formation, which is an essential step involved in blood clotting. But from a visual observation we found that chitosan enhances wound healing in diabetic condition. Our result also showed that chitosan cannot normalize platelet aggregation in diabetic conditions. These two results point out that there is no relation between chitosan-induced platelet aggregation and wound healing. Biagini *et al.*[16]. reported that in patients undergoing plastic surgery, the donor sites were treated with *N*-carboxybutyl chitosan, which enhanced wound healing. This *N*-carboxybutyl chitosan leads to formation of regularly organized cutaneous tissue and reduces anomalous healing. Although chitosan was not found to help in normalizing the altered platelet aggregation, it helps in wound healing. Thus, chitosan-induced enhanced wound healing in diabetes observed in the present study may be due to organization of newly formed tissue and thus help in damage repair.

In conclusion this study presents the effect of the bioactive polymer chitosan in wound healing in diabetes mellitus.

1. Ditzel, J. and Rooth, G., *Diabetes*, 1955, **4**, 474–478.
2. Horne, W. C., Anderson, G. M. and Cohen, D. J., *Am. J. Hematol.*, 1991, **38**, 48–53.
3. Olsen, R., Schwartzmiller, D., Wepner, W. and Winandy, R., in *Chitin and Chitosan* (ed. Anthonsen, T. and Sandford, P.), Elsevier publ., 1989, I edn, pp. 813–819.
4. Makino, N., Dhalla, K. S., Elimban, V. and Dhalla, N. S., *Am. J. Physiol.*, 1987, **253**, 202–207.
5. Domino, S. E., Repaske, M. G., Bonner, A., Kennedy, M. E., Wilson, A. L., Brandon, S. and Limbird, L. E., in *Methods Enzymol.*, 1992, **215**, 181–200.
6. Frojmovic, M. M., Milton, T. G. and Duchstel, A., *J. Lab. Clin. Med.*, 1983, **101**, 964–976.
7. Jamaluddin, M. P. and Krishnan, L. K., *J. Biochem. Biophys. Methods*, 1987, **14**, 191–200.
8. Junod, A., Lambert, A. E., Stauffacher, W. and Renold, A. E., *J. Clin. Invest.*, 1969, **48**, 2129–2139.
9. Born, G. V. R., *Nature*, 1962, **194**, 927–929.
10. Hantgan, R. R., *Blood*, 1984, **64**, 896–906.
11. Kwaan, H. C., Colwell, J. A., Cruz, S., Suwanwela, N. and Dobbie, J. G., *J. Lab. Clin. Med.*, 1972, **80**, 236–239.
12. Takiguchi, Y., Wada, K., Matsuno, H. and Nakashima, M., *Thromb. Res.*, 1991, **63**, 445–456.
13. Dunbar, J. C., Reinholt, L., Henry, R. L., Mammen, E., *Diabetes Res. Clin. Pract.*, 1990, **9**, 265–272.
14. Toth, L., Szenasi, P., Jambor, G., Kammerer, L. and Romics, L., *Diabetes Res. Clin. Pract.*, 1992, **2**, 143–148.
15. Zucker, M. B., *Sci. Am.*, 1980, **242**, 86–103.
16. Biagini, G., Bertani, A., Muzzaarelli, R., Domadei, A., DiBenedelto, G., Belligolli, A., Riccotti, G., Zucchinic, R. C., *Biomaterials*, 1991, **3**, 281–286.

ACKNOWLEDGEMENTS This work was supported by a grant from Department of Atomic Energy, BARC, India. PSP thanks CSIR for a fellowship.

Received 19 June 1995; revised accepted 15 July 1995

Permeability of cuticle of *Helicoverpa armigera* (Hübner) larvae to deltamethrin

A. P. Padma Kumari, A. Phokela and K. N. Mehrotra

Division of Entomology, Indian Agricultural Research Institute, New Delhi 110 012, India

Permeability of the cuticle of *H. armigera* larvae to deltamethrin has been studied in three populations originating from Delhi, Sirsa and ICRISAT. Delayed penetration of topically applied ^{14}C-deltamethrin was observed in the Sirsa population. This suggests that the permeability of the cuticle plays an important role in imparting resistance in the Sirsa population, whereas in Delhi and ICRISAT populations it is not very important.

REDUCED permeability of the cuticle may be a factor conferring resistance in insect pests to insecticides[1-6]. *Helicoverpa armigera* (Hübner) has acquired resistance to synthetic pyrethroids (SPs) in major parts of India[7-10]. Recently, by a very indirect method some evidence was presented to suggest that the cuticle may, perhaps, play some role in manifestation of resistance to SPs in *H.*

armigera populations of Australia[11]. This led us to reinvestigate the permeability of cuticle to *H. armigera* larvae obtained from three widely separated areas, Sirsa (Haryana), Delhi, and ICRISAT, Andhra Pradesh. We report here that the reduced permeability of the cuticle plays a significant role in imparting resistance to *H. armigera* populations originating from Sirsa.

Final instar larvae of *H. armigera* used in the present study were collected from chickpea fields in Delhi, from cotton fields in Sirsa (village Balasar) on cotton and from weeds at ICRISAT. The Sirsa population had been subjected to local practices of plant protection prevalent in that area. But all the populations were conditioned on artificial diet[10]. For studying permeability, 0.1 µg ^{14}C-deltamethrin, about 80,000 Bq in one microlitre of methanol, was applied on the thoracic dorsum of the final instar larvae (weighing 350–375 mg), using an electrically operated microapplicator. Larvae treated for different time intervals were held in scintillation counting vials without any food. At intervals of 10 min, three larvae were washed individually with 5 ml of the scintillation mixture (0.55% w/v 2,5-diphenyl oxazole in 2 : 1 toluene and methyl cellosolve). These washings, having the remaining deltamethrin on cuticle, were estimated in a liquid scintillation counter[12]. The percentage of the remaining activity on the thoracic dorsum when plotted (on log scale) against time gave a straight line, suggesting that deltamethrin followed a first-order kinetics in permeating the cuticle. The $t_{0.5}$ value was computed for all the three populations.

The permeability of cuticles in Delhi and ICRISAT populations did not differ and was almost the same as that observed earlier[10]. However, the permeability was significantly different in populations originating from Sirsa. The computed $t_{0.5}$ value was 42.00 min for the Delhi population and 39.96 min for the ICRISAT population. In contrast to this, $t_{0.5}$ was 71.91 min for the Sirsa population (Table 1). This indicated that ^{14}C-deltamethrin had taken almost twice the time to penetrate the cuticle of the Sirsa population than the other two. The permeability of the cuticle of all the three populations was observed to be monophasic and followed first-order kinetics (Figure 1), which was similar to that reported earlier for insect cuticle and animal integuments[13–15]. It may also be mentioned that in a set of experiments on

Table 1. Permeability of the cuticle of final instar larvae of *H. armigera*

Location	$t_{0.5}$ (min)	Regression equation	r^2
Delhi	42.00	$Y = 2.0 - 0.00716x$	0.9445
ICRISAT	39.96	$Y = 2.0 - 0.00754x$	0.6789
Sirsa	71.91	$Y = 2.0 - 0.00419x$	0.9456

Note: $t_{0.5}$ is the time required for penetration of 50% of the applied dose. All values were significant at 5% level.

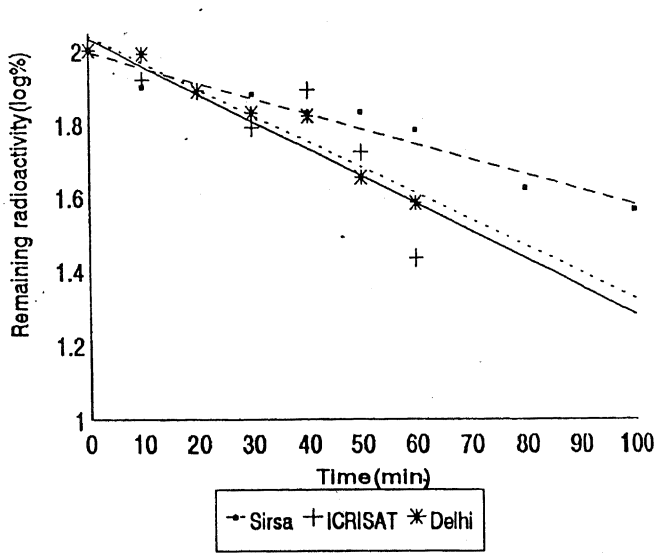


Figure 1. Permeability of *H. armigera* larval cuticle to ^{14}C-deltamethrin.

the metabolism of ^{14}C-deltamethrin it was found that metabolism was high in the microsomal fraction of cuticular homogenates of the Sirsa population than in the same of the Delhi population, the ratio being about 2 : 1 between Sirsa and Delhi. These results, therefore, suggest that poor permeability of deltamethrin along with high rate of metabolism in the cuticle itself was important in imparting resistance to the Sirsa population. It is likely that *H. armigera* populations have developed resistance to SPs by different mechanisms in different geographical/agro-ecological zones of the country. It is also possible that in the Sirsa population the cuticle is an important factor contributing towards resistance. Thus, the present studies suggest that 'PEN' mechanism, as suggested by Gunning *et al.*[16] in Australian strains, may play a prominent role in imparting resistance to SPs in the Sirsa population.

1. Chang, C. K. and Jordan, T. W., *Pesticide Biochem. Physiol.*, 1982, **17**, 196–204.
2. Delorme, R., Fournier, D., Chaufaux, J., Cuany, A., Bride, J. M., Auge, D. and Berge, J. B., *Pesticide Biochem. Physiol.*, 1989, **33**, 37–45.
3. Golenda, G. F. and Forgash, A. J., *Pesticide Biochem. Physiol.*, 1989, **33**, 37–45.
4. Noppun, V., Saito, T. and Miyata, T., *Pesticide Biochem. Physiol.*, 1989, **33**, 83–87.
5. Little, E. J., McCaffery, A. R., Walker, C. H. and Parker, T., *Pesticide Biochem. Physiol.*, 1989, **34**, 58–68.
6. Rees, H. H., *Insect Biochemistry*, Chapman and Hall, 1977, p. 63.
7. Dhingra, S., Phokela, A. and Mehrotra, K. N., *Natl. Acad. Sci. Lett.*, 1988, **11**, 123–125.
8. McCaffery, A. R., Maruf, G. M., Walker, A. J. and Styles, K., Brighton Crop Protection Conference, 1988, vol. 4C-14, pp. 433–438.
9. McCaffery, A. R., King, A. B. S., Walker, A. J. and El-Nayir, H., *Pesticide Sci.*, 1989, **32**, 383–384.

10. Phokela, A. and Mehrotra, K. N., *Proc. Indian Natn. Sci. Acad.*, 1989, **B55**, 235–238.
11. Kennaugh, L., Pearce, D., Daly, J. C. and Hobb, A. A., *Pesticide Biochem. Physiol.*, 1993, **43**, 234–241.
12. Ishaaya, I. and Casida, J. E., *Pesticide Biochem. Physiol.*, 1980, **14**, 173–184.
13. Buerger, A. A., *J. Theor. Biol.*, 1966, **11**, 131–135.
14. Buerger, A. A., *J. Theor. Biol.*, 1967, **14**, 66–73.
15. Hadgraft, J. and Brian, K. R., *Pesticide Sci.*, 1990, **30**, 81–89.
16. Gunning, R. V., Easton, C. S., Balfe, M. E. and Ferris, I. G. *Pesticide Sci.*, 1991, **33**, 437–490.

ACKNOWLEDGEMENTS. We thank Roussel Uclaf, Procida, France, for providing labelled deltamethrin. This article is a part of the Ph D thesis submitted by the senior author to the Post-Graduate School, IARI, New Delhi.

Received 3 August 1994; revised accepted 14 July 1995

The 'basic' vs the 'evolved'

Quark and Jaguar. M. Gel-Mann. Little Brown and Co., London. 1994. 392 pp. Price: £ 18.99.

That Murray Gell-Mann's latest book gives the pride of place to the quark as the most basic building block of matter will presumably not be contested by its less dramatic partner, the lepton, since it must be 'aware' of the 'creator', in company with another genius[1], of this bizarre yet most concrete of all objects, especially after the announcement of the last and most elusive member[2] of the series. However, the corresponding privilege for the jaguar to qualify as a prototype for the most complex adaptive system on earth is more a matter of personal taste. (The choice could well have fallen on, say, the giraffe[3] if it were a different author). In any case one cannot but feel a sense of admiration for the poet (Arthur Sze) who had both the words (the quark and the jaguar) in his poetic dictionary and thus captured the literary fancy of the great physicist. This is somewhat reminiscent of an earlier occasion (three decades ago) when James Joyce, through his *Finnegans Wake*[4], had (inadvertently?) supplied the 'name' of the first member of this book.

While the title of a book is largely the prerogative of the author, the more relevant issue is the subject matter and the language in which it is intended to be conveyed. The general tone and the tenor of the book seems to be in keeping with the trend set by some of the author's distinguished contemporaries (Feynman and Dyson) to write personal glimpses on physics in an 'autobiographical' style, but since the author himself disclaims any autobiographical intentions, one must look for other, say educational, motivations designed to convey to the lay reader the significance of major discoveries in a frontline field in a sufficiently nontechnical language without too much sacrifice of rigour. Even such ventures are no longer considered novel, since after the debut of the *Tao of Physics*[5] nearly two decades ago, many impressive literary pieces dealing with abstract concepts in physical and mathematical sciences in a nontechnical manner by professional scientists of great distinction have appeared on the scene[6,7]. And now Gell-Mann has joined this select club with his own 'coup de grace'. In view of the great stature of the author, this book is bound to attract a good deal of attention, as well as an *inter-se* comparison with some of the classics[6,7] on cosmology on the one hand and other equally authoritative treatises[8] on a major aspect (chaos) of its theme on the other. Without presuming to pass a hasty judgement on the book, the issue of immediate concern is its 'thrust', together with the nature of the emphasis that the book wishes to convey.

What is the main thrust of the book? Fortunately, the distinguished author has himself provided enough summary on his own (at both ends of the book) to obviate any need on the reader's part to make an independent inference in this regard. The book is written in an intensely personal style and is studded with examples taken from the author's own experiences from family and friends alike. The entire setting is centred around the Santa Fe Institute (SFI) of Interdisciplinary Studies, which the author founded about a decade ago (after taking an early retirement from his home institution – Caltech) and which now represents his principal scientific base. Shorn of such personal details, the book is presented in four parts, in an obvious attempt to address a whole gamut of information, from 'the simple: a quark inside an atom' to 'the complex: a jaguar prowling its jungle territory in the night', together with the innumerable interrelationships among the various components that make up the whole and are continually showing dynamical manifestations in an ever-growing evolutionary scenario.

The first part of the book is an attempt to describe the relationships between the concepts of simplicity and complexity, especially of *complex adaptive systems* (CAS) – child learning a language, bacteria developing resistance against antibiotics, human scientific enterprise, and so on. The role of theory in science is discussed, 'as well as the issue of which sciences are more fundamental than others', along with the related question of reductionism. The second part deals with the fundamental laws of physics, those governing the cosmos and the elementary particles, out of which all matter in the universe is composed. Here the quark comes into its own (as does its creator); so do superstrings, which the author believes 'for the first time in history' to 'offer the serious possibility of a unified theory of all particles and forces of nature'. The third part is devoted to the various selection pressures operating in complex adaptive systems, 'especially in biological evolution, human creative thinking, critical and superstitious thinking' (which he strongly condemns), and some aspects (including economic ones) of the behaviour of human societies. Many *ad hoc* concepts, such as fitness and fitness landscapes are introduced along the way. The fourth and final part is concerned with policy matters (not science) and 'with advocacy'. (This reviewer hopes that the last phrase does not imply that the author considers science to be so objective as to be above policy and advocacy, since science does seem often to exhibit a high degree of subjectivity at the *operational* level)[9,10].

Now to some essential details concerning the different parts of the book. The most serious discussion (understandably) concerns the *second* part, which bears the characteristic imprint of the author's seminal contributions to elementary particle physics during the fifties and sixties (including the quark proposal among other things), as well as some fresh 'hard physics' on the foundations of quantum mechanics (carried out at his SFI). Since this topic dealing with the deeper issues of 'measurement' and 'paradoxes' is relatively new and still confined to the pages of technical journals, its wider appeal to the scientific community at large warrants some elaboration in this review (see below). However, the other, more familiar topics of the second part, as well as some from the first part, are treated rather briefly (within the limited space available), while Parts 3 and 4 are too 'righteous' (and personalized) for a cogent review.

Complex adaptive systems

Much of the discussion in Part 1 is taken up with a discursive analysis of complex adaptive systems (CAS), their patterns of evolution by acquiring information about their respective environments through their own interactions with the latter, identifying regularities in the information, condensing the (abstracted) regularities into a kind of schema or model, and acting in the real world on the basis of

that schema. Their capacity to learn from experience is itself a product of biological evolution, which in turn gives rise to other new types of CAS (e.g. immune system in mammals). Indeed, a CAS generally has a tendency to generate other such systems in turn. They are, of course, subject to the laws of nature, which themselves rest on the fundamental physical laws of matter and the universe. And (over and above the quantum-mechanical (q.m.) effects) indeterminacies arise, even in classical motion, partly due to an *approximate* knowledge of the initial conditions and partly due to *measurement imperfections*. The latter give rise to CHAOS[8], which tends to amplify the resultant indeterminacies (whether from q.m. effects and/or from measurement imperfections) to macroscopic levels. The entire discussion is studded with a rich vocabulary of terms with subtle distinctions in their meaning, depending upon the context of their use. Different types of complexity, along with several associated concepts (conciseness, crudity, algorithmic information content, information, compression, randomness, effective complexity, direct adaption vs CAS) abound in the text.

In Gell-Mann's view, the concept of CAS is well illustrated by the human scientific enterprise. The schemata are theories, and what takes place in the real world is the confrontation between theory and observations. This is strongly reminiscent of Holton's[11] definition of themata along with the associated concept of the contingent plane, which Holton had developed very systematically in the early seventies. (A summary of the main ideas involved may be found in ref. 10). In this regard Gell-Mann strongly endorses the dictum of philosopher Karl Popper concerning the falsifiability (by experimentation) of scientific theories. His views on the scientific enterprise as a self-correcting process (despite occasional abuses) as well as his endorsement of the themes of simplicity and unification in physical theories are quite standard, but his joining the issue, on grounds of semantics, with a distinguished contemporary thinker, Thomas Kuhn on his seminal theme of paradigm shifts[9] does not seem to carry the same degree of conviction.

Gell-Mann goes into great lengths in analysing the powers of theory at successively deeper levels of operation, stressing in the process its unifying role in bringing together apparently unrelated phenomena under a more transparent pattern which reveals their relationships, thus facilitating a more 'compressed' (compact?) description. He illustrates the idea with the simple example of how a relatively empirical model like Zipf's law has eventually evolved in the hands of deep thinkers (like Mandelbrot) into something quite profound: fractal behaviour and scale independence. (One may add in parentheses that it already has the germs of renormalization group theory). In this context he introduces several related terms like depth and crypticity as a sort of figure of merit for theorizing.

The quantum universe

Coming to the fundamental laws of matter and the universe, the distinguished physicist is at last on his own territory where he has every reason to look back with satisfaction at the culmination of his quark proposal[1,4] into the standard model of today. For good reasons he does not obviously believe in the finality of the so-called grand unified theories. He also does not have good words to say about Einstein's mental faculties during his twilight years of research on a unified field theory, pointing out several specific flaws in his approach due to the neglect of fields (weak and strong) other than the gravitational and electromagnetic as well as fermion fields, and, most importantly, the need to construct a unified theory within the framework of quantum mechanics. On the other hand, he seems to believe that Einstein's dream may at last have been realized with the superstring theory, which in his view has the right ingredients to accomplish the unification, since it predicts in a natural way the existence of the graviton within its broad framework. In particular, the heterotic superstring theory, which realizes this crucial quantum in a (conceptually) convincing fashion seems to fit the bill quite well without running into the perennial problems of infinities. Reasons? (i) It explains the great multiplicity of particles discovered and yet to be discovered in the laboratory. (ii) It does not contain any arbitrary particles or forces. (iii) It emerges from a simple and beautiful principle of self-consistency (which dates back to the bootstrap idea of the sixties just preceding the quark model).

Even if the superstring turns out to be eventually 'correct', its claims to TOE (theory of everything) is still beset with other problems. For, the actual evolution of the universe depends not only on the physical laws but also on the initial conditions plus the unknown effects of quantum fluctuations, which in turn are closely linked with the 'arrow of time' (On this topic, a nontechnical yet highly readable description may be found in ref. 6). Regarding the initial conditions, Gell-Mann recognizes that a good candidate is one already suggested by the Hawking–Hartle theory[6], viz. the no-boundary condition, which, along with the (weak) anthropic principle, is apparently able to account for the arrow of time[6]. On the other hand, Gell-Mann himself believes that since the laws of physics are quantum-mechanical (not deterministic) and permit only the calculation of probabilities of various alternative histories of the universe that describe different ways events could play themselves out (given the initial conditions), information about which of those sequences of events is actually occurring can be gathered only from observation and is supplementary to the basic laws themselves. And further, each alternative history of the universe depends on the results of an inconceivably large number of accidents. The chance outcomes of the latter (*á la* quantum mechanics) have 'determined' almost every aspect of evolution – from the universe in general to the various biological species on earth. Thus, the algorithmic information content of each alternative history of the universe merely gets a tiny contribution from the 'simple' basic laws, but a gigantic contribution from all the 'quantum accidents' that arise along the way!

A modern view of quantum mechanics

Some well-known paradoxes in quantum mechanics – the Schrödinger cat and the EPRB paradox in particular – have generated much controversy and learned discourse in the recent literature. In this book a new perspective is offered on these and related issues on the basis of Hugh Everett's theory of wave function of the entire universe, which has since been explored by many authors, including the famous Hawking–Hartle formulation with the same title. Gell-Mann and Hartle

have since investigated the foundations of quantum mechanics on these lines with a view to clarifying the working of the latter in the quasiclassical domain. Briefly, the picture which makes free use of the notion that *a given system can have different possible histories, each with its own probability*, is as follows: Suppose a set of alternative histories of the universe is specified, and that these histories are exhaustive and mutually exclusive. Rather paradoxically, quantum mechanics *does not always* assign a probability to each one. For a pair of such histories a necessary condition for a meaningful assignment of their respective probabilities is that they *do not interfere* with each other.

What is the condition of noninterference? The more fine-grained the history is, the more are the chances of interference. Therefore, in order to have actual probabilities it is necessary to consider histories that are sufficiently *coarse-grained*, since coarse-graining can wash out interference terms. In particular, if the interference term between each pair of coarse-grained histories is zero to a very good approximation, then all the coarse-grained histories are said to *decohere*, and one gets true probabilities.

But why does decoherence occur? It is the 'entanglement' of what is followed in the coarse-grained histories with 'what is ignored or summed over' that holds the key to the said mechanism. This does not give anything effectively new in the classical or quasi-classical domains, but helps solve several standard puzzles where quantum effects are apparently important. Thus, in the famous example of the Schrödinger cat, the alleged quantum interference between the live and dead cat scenarios can be explained away by the fact that there is plenty of opportunity for decoherence between the coarse-grained histories in which the cat lives and the corresponding histories in which the cat dies. Thus, the *live and dead cat scenarios decohere*, hence no interference.

By and large a very similar mechanism is involved in the Einstein–Podolsky Rosen–Bohm (EPRB) experiment concerning the decay of a spinless particle at rest into two photons wherein the state of polarization (circular or plane) of one is supposed to 'determine' that of the other. Now quantum mechanics says that both linear and circular polarizations cannot be simultaneously specified (these alternatives 'sit' on different branches of history, and cannot be considered together), while Einstein's 'completeness' would imply simultaneous assignments for both, perhaps through some 'hidden variable' mechanism. To distinguish between these two possibilities, Bell[12] had found some testable inequalities for the correlation function of the two photon polarizations wherein the latter could attain values in quantum mechanics that are not allowed in a classical hidden variable theory.

Subsequent experimentation (by Allan Aspect) confirmed the quantum-mechanical result, but led to a flurry of speculations about quantum mechanics showing bizarre properties, the implication being that measuring the polarization (circular or plane) of one photon somehow affects the other photon. On the other hand, there is nothing mysterious in all this according to the logic of decoherence given above, since the measurement process *does not* cause any physical effect to propagate from one photon to the other. Thus, if on a particular branch of history, the plane of polarization of the one photon is measured (and thereby specified with certainty), then on the *same branch* of history the plane polarization of the other photon is also specified with certainty. On a different branch of history, the circular polarization of one of the photons may be measured, in which case again the circular polarization of the other photon is also specified with certainty. This *does not* imply any *signal* from one photon to the other in the actual experiment; hence, no action at a distance is involved despite the seemingly instantaneous effect, implying faster-than-light propagation. Indeed, the physicist in the author comes down heavily on such distortions in the literature.

Summary and conclusion

As for the third and fourth parts of the book, no attempt is made at further abstraction, except to draw some conclusions from the author's own 'afterword' (Chapter 23): The book is *not* a treatise, but a running commentary on the various multidisciplinary researches in progress at the SFI. Running through the entire text is an essential 'interplay' between the fundamental laws of nature and the operation of chance, since the fundamental laws are quantum-mechanical, and quantum mechanics supplies only probabilities for alternative coarse-grained histories of the universe. This often results in their branching, with related probabilities. On certain branches and at certain times and places in the universe the conditions are ripe for the evolution of CASs which function best in a regime intermediate between order and disorder. They exploit the regularities provided by the approximate determinism of the quasiclassical domain, and at the same time they profit from the indeterminacies in their quest for 'better' schemata.

A CAS discovers 'regularities' in its incoming data stream by noticing features in common with its own, and emerges with a higher form of complexity through self-organization. When a CAS gives to a new kind of CAS, whether by aggregation or otherwise, the same can be considered a *gateway event*, e.g. the evolution of the mammalian immune system. And finally, many such features of biological evolution are found in other CASs such as human thought, social evolution and adaptive computing. All these systems keep exploring possibilities, opening up new paths, discovering gateways, and occasionally giving rise to new types of CAS. And the process goes on and on.

Looking back on this broad-spectrum book, one wonders as to what kind of clientele it is intended for, since, apart from the contents of Part 2, the rest of the book, by the very nature of its scope, is too full of definitions and *ad hoc* concepts to do 'coherent justice' to the vast canvas of material that the author has taken upon himself to present to the readership. This is certainly an unenviable task, for its theme is much more diffuse and tentative (many of the conclusions, in the author's own phrase, are the results of ongoing activities which are not yet firmly established) than e.g. the themes of other such books[6,7], which are more focused and based on firmer investigations. On the other hand, since the book is written in a nontechnical style, it must be assumed that it is meant for the widest possible audience, captive or otherwise. While the captive audience may be already aware of its contents along with the various jargons employed, one is not so sure if the general reader will find it equally enjoyable, despite the absence of technicalities. Moreover, the repeated emphasis on SFI and frequent references to the work of only a few personally known to the author (irrespective of the

depth of their investigations) does not quite make for a smooth reading from an educative angle, unless the reader were warned in advance that he was reading an autobiography.

Gell-Mann has long been my hero in physics, perhaps next only to Feynman and Dyson. I had richer expectations from the entire book but was only partially rewarded with Part 2 of it.

1. Zweig, Z. G., CERN-8182/TH-401, 1963, unpublished.
2. C. D. F. Collaboration, Fermilab, 1994–95; also D-O Collaboration, 1995.
3. Davies, P., in *Doing Science* (Reality Club) (ed. Brockman, J.), Prentice-Hall, New York, 1991.
4. Gell-Mann, M., *Phys. Lett.*, 1964, **8**, 214.
5. Capra, F., *The Tao of Physics*, Flamingo, London, 1976.
6. Hawking, S., *A Brief History of Time*, Bantam, New York, 1988.
7. Weinberg, Steven, *The First Three Minutes*, Flamingo, London, 1977.
8. Prigogine, I. *et al.*, *From Order to Chaos*,
9. Kuhn, T., *Structure of Scientific Revolutions.*, University of Chicago Press, Chicago, 1970.
10. Mitra, A. N., *Curr. Sci.*, 1993, **64**, 162.
11. Holton, G., *From Kepler to Einstein*, Harvard University Press, Harvard, 1971.
12. Bell, J., *Physics*, 1964, **1**, 195–200.

A. N. MITRA

National Institute of Advanced Studies, Indian Institute of Science Campus, Bangalore 560 012, India.

Functional Dynamics of Phytophagous Insects. T. N. Ananthakrishnan, ed. Oxford & IBH, 66, Janpath, New Delhi 110 001. 1994. viii + 304 pp. Price: Rs 550.

One-third of a million insect species are known to be phytophagous and constitute one-quarter of all living beings[1]. The number of recorded lepidopteran species alone, which are mostly phytophagous, is one order of magnitude greater than all species of birds and mammals put together – a staggering diversity which continues to overwhelm entomologists. The obvious success of phytophagy as a mode of life for insects testifies to their triumph in the evolutionary battle against plants. But plants have not been silent victims of insect onslaught either. They have waged their own evolutionary battle and what one sees today is a dynamic equilibrium in a continuing race. Interactions between insects and their host plants were first put into a comprehensive evolutionary perspective by Ehrlich and Raven's[2] classic paper, which triggered off an intense flurry of research activity. So strong has been its influence that almost all of plant morphology and plant secondary chemistry have since been attributed to selective pressure exerted by phytophagous insects. The book under review, coming close on the heels of another landmark publication in this area[3], is a laudable attempt at summarizing the mind-boggling literature on ecological and evolutionary aspects of insect–plant interactions.

Is geographic variation in plant allelochemicals of any significance to insect herbivores? In their chapter Johnson and Scriber address this question with examples from nonagricultural systems by dissecting out the ecogeographic variation in plant chemistry into historical, genetic, altitudinal and latitudinal factors. And how do phytochemicals alter insect behaviour? The chapter by Norris gives a brief overview of the major classes of chemical messengers, which is followed by a detailed discussion on the mechanism by which insects respond to chemicals. The idea that many common chemical messengers are both elicited and perceived by sulphydryl/disulphide-dependent mechanism of the receptors is convincingly argued. His call for detailed investigations of chemical communication at electrochemical level for a better understanding of cell–cell and environment–cell exchange of chemically based energy would hopefully attract more chemists to biology. The chapter on the physiological basis of feeding and oviposition behaviour in moths by Ramaswamy is an excellent review of the current knowledge on the role of proximate factors in host finding in moths. At the simplest level, host recognition and acceptance involves integration of inputs from several sensory modalities such as vision, mechanoreception and chemoreception. The internal metabolic state of the insect, the ontogenetic and phylogenetic influences on the mechanism of host selection add several levels of complexity to the dynamics of insect–plant interactions.

Do plants possess chemicals that provide both positive and negative stimuli and influence host selection? Dethier[4] had suggested that a balance of positive and negative stimuli could be present in the plants and the ultimate behavioural response of the insect depends on the balance between internal state of the insect as regards motivation and degree of satiety and external stimuli. This was further elaborated by Miller and Strickler[5], who proposed a 'rolling fulcrum model' wherein the internal excitatory/depressor factors could tip the balance in favour of host acceptance or rejection. This issue is examined by Renewick and Huang in the chapter on chemical stimuli mediating oviposition by lepidoptera. The authors review their work with the *Pieris–Crucifer* system, wherein the host plants have chemicals that are both oviposition stimulants and deterrents. Their results show quite convincingly that final rejection or acceptance of a host depends on the insect's assessment of the balance of stimulants and deterrents.

Ananthakrishnan, Venkatesan and Sridhar examine the effect of relative levels of nutrients and other chemicals like phenols, fatty acids and flavonoids in different plant parts on feeding and reproduction in the lepidopteran pest complex of cotton. The authors argue that this information can be of much use in managing insect pests of cotton by manipulating plant chemistry. Uthamasamy reviews the behavioural processes involved in host selection by the bollworm complex on different varieties of cotton. The emphasis here, however, is more on the effect of plant morphology on host acceptance by bollworms.

'Your enemy's enemy is your best friend.' That even plants believe in this is convincingly shown by Whitman and Nordlund in the seventh chapter. While reviewing the tritrophic interactions between plants, herbivores and their natural enemies, the authors present evidence to show that plants actually 'communicate' with natural enemies of their enemies, i.e. with the enemies of herbivores. Damage by herbivores induces the plants to produce chemicals that not only attract but actually serve as road maps to the natural enemies of herbivores. The chapter ends with a tempting speculation that these chemical beacons, 'kairomones' as they are called, hold great potential for solving pest problems.

Whitaker, Blum and Slansky Jr. attempt to trace the possible evolutionary path of

carnivory in phytophagous insects. The authors appeal for more careful and detailed study of such facultative carnivory since many instances of carnivory are either overlooked or not reported.

Another insightful and particularly enjoyable chapter also tackles the same theme, viz. plant defence against herbivores through tritrophic interaction. Whitman takes us on a fascinating journey into the world of plants and the elaborate strategies (nonchemical) employed by them to defend themselves against herbivores. The role of floral nectar, pollen, extrafloral nectaries, food bodies for ants, and domatia (shelters) in plant defence against herbivores by encouraging action of natural enemies is discussed in detail.

Interaction between gall-inducing insects and their host plants is reviewed by Raman in perhaps the toughest chapter, if one were to consider the extent of information published. The burgeoning literature on adaptional integration of host plant and galling insect species, host chemistry, chemical ecology of host selection, insect feeding, gall induction and the response of the host tissue is effectively summarized, but perhaps not sufficiently synthesized.

When two organisms are so closely linked as a plant and its insect pest/guest, the extinction of one is bound to increase the chances of extinction of the other. The issue of host specificity, metapopulations and conservation, using *Drosophila magnaquinaria* and its host plant as the model system, is addressed in the last chapter and the impact of small patches of host plants on the populations of the monophagous fruit fly is discussed.

All the chapters of this book, written by experts in their respective fields, are well researched. A more or less uniform presentation of chapters is also commendable for a multiauthored book as this one. The book, however, has its share of printer's devils. For instance, 'Lepidoptera' is spelt wrongly in the contents page and the legend to Figures on p. 253 is incorrect. Although each chapter can be read independently, they could still have been arranged into smaller subsections following a theme. Overall, the book offers a wealth of information to entomologists and will serve as an excellent companion volume to the recently published book on a related theme[3]. Undoubtedly, it should find a place on every library shelf. But unless publishers in India produce inexpensive paperback editions to accompany hard-cover library editions, the habit of buying books for personal use will surely go extinct.

1. Strong, D. R., Lawton, J. H. and Southwood, T. R. E., *Insects on Plants: Community Patterns and Mechanisms*, Blackwell Scientific Publications, London, 1983, pp. 308.
2. Ehrlich, P. R. and Raven, P. H., *Evolution*, 1964, **18**, 586–608.
3. Rosenthal, G. A. and Berenbaum, M. R., *Herbivores: Their Interactions with Secondary Plant Metabolites*, Academic Press, New York, 1992, vols I & II, pp. 488 and pp. 493.
4. Dethier, V. G., *Entomol. Exp. Appl.*, 1982, **31**, 49–56.
5. Miller, J. R. and Strickler, K. L., in *Chemical Ecology of Insects* (eds Bell, W. and Carde, R), Chapman and Hall, London, 1984, pp. 127–155.

K. CHANDRASHEKARA
RAGHAVENDRA GADAGKAR*,†

Department of Entomology,
University of Agricultural Sciences,
GKVK, Bangalore 560 065, India.
**Centre for Ecological Sciences,*
Indian Institute of Science,
Bangalore 560 012, India.
†Jawaharlal Nehru Centre for
Advanced Scientific Research,
Jakkur P.O., Bangalore 560 064, India.

The State of Food and Agriculture 1993. FAO Agriculture Series No. 26, Food and Agriculture Organization of the United Nations, Rome. 1993. 306pp.

The State of Food and Agriculture 1994. FAO Agriculture Series No. 27, Food and Agriculture Organization of the United Nations, Rome. 1994. 357pp.

The Food and Agriculture Organization of the United Nations (FAO), established in October 1945 for helping to achieve freedom from hunger and to disseminate widely data and information relevant to this basic goal, completes 50 years of service this year. Since its establishment, FAO has been publishing a wide range of journals, reports and books. Invariably, these are the most authentic sources of data available in the world concerning area, yield, production, prices, imports and exports, and other relevant information relating to agriculture, forestry and fisheries. FAO has also an early warning system to draw attention to hot spots in the field of agricultural production so as to help in taking proactive steps to avoid famines. Early warning systems also exist in the case of pest epidemics. A good example is the locust warning system. As it celebrates its 50th Anniversary, FAO can look back with satisfaction on the varied services it has rendered to member nations and more particularly in the field of communication and publication. Most of the FAO publications, including the books under review, are now available as electronic products.

The books under review bring together in a very readable form the significant developments in agriculture during the preceding years. The reports are structured in a way that they contain generic information as well as specific data relating to regions and countries. In addition, one major issue is taken up for detailed analysis in each of the reports. For example, the 1993 report deals with water policies and agriculture, while the 1994 report pays specific attention to various development and policy dilemmas in the forestry sector. The reports give the reader in a capsule form the situation on the food front. FAO's definition of agriculture includes all land- and water-based occupations such as crop husbandry, animal husbandry, fisheries and forestry. Both production and trade issues are considered. The economic environment under which agriculture progresses is often dealt with in an objective manner.

The special sections on water and forests in the 1993 and 1994 reports have one point in common. They both emphasize the need for community participation in the sustainable management of these two essential foundations for sustainable advances in agricultural production. Unless there is equity in sharing the benefits, the requisite degree of public cooperation in managing these assets will not be forthcoming. Participatory water and forest management involves a paradigm shift in management culture, moving away from an entirely government-controlled management to the involvement of all the stake holders in both decision-making and implementation mechanisms.

On the occasion of the 50th anniversary of FAO, the question can be asked as

to why endemic hunger arising from undernutrition and malnutrition is still widely prevalent, although the founding fathers of FAO had felt even in 1945 that the world has the requisite technical and financial resources to make hunger a problem of the past. Is this situation due merely to a lack of the requisite will at the national and global levels to give operational content to the political slogan of 'food for all'?

On the basis of a recent study, the International Food Policy Research Institute (IFPRI), located in Washington, DC, USA, believes that by the year 2020 significant reductions can be made in poverty, food insecurity, malnutrition and micronutrient deficiencies and significant advances made toward efficient, effective and low-cost food and agricultural systems that are in harmony with the natural resource base. The fact remains that although today the world produces enough food to feed each and every person, nearly 800 million children, women and men remain food-insecure. The reasons for this tragic irony are to be found in the nonfood sector, such as lack of purchasing power and entitlements, poor sanitation and health care and non-availability of clean drinking water. It is, therefore, time that we redefine our concept of food security as '*livelihood security for the households and all members within, which ensures physical and economic access to balanced diets, including the needed micronutrients, safe drinking water, environmental sanitation, basic health care and primary education*'. Such a redefinition will help to articulate the public policy package needed to ensure food security at the level of each individual in a household. The emphasis on individuals rather than on families is important, since in many societies women and girl children tend to suffer more from undernutrition and malnutrition and there is a trend towards an increasing feminization of poverty.

How do we achieve this goal? In October 1995, a meeting of representatives of Member-Nations of FAO will be held at Quebec in Canada to commemorate the signing FAO's Charter in that city in October 1945. This is an appropriate occasion to analyse why the Freedom from Hunger Campaign launched in the early sixties by the late Dr B. R. Sen when he was the Director-General of FAO is still to achieve its objectives. The 50th Anniversary of FAO is best commemorated by the adoption of a *Global Convention on Sustainable Food and Livelihood Security*. Such a convention can include the following major provisions:

1. Implement both in letter and spirit the plans of action agreed to at recent UN Conferences, particularly at Rio, Cairo and Copenhagen and take speedy action to implement the provisions of the Desertification Convention.

2. Ensure adequate support for research and development efforts designed to promote national and international public good, since the fast-spreading Intellectual Property Rights (IPR) environment is likely to promote greater investment only in short-gestation-period and profit-driven research. Integrating the dimension of sustainability in research objectives, however, involves location-specific and participatory research with rural families. This kind of research is unlikely to attract adequate investment from the corporate sector.

3. Ensure the rapid dissemination of environmentally sound agricultural technologies, particularly among resource-poor farm families, who may have difficulties in gaining access to green technologies as a result of the provisions of the Trade-related Intellectual Property Rights (TRIPS) of the World Trade Agreement.

4. Promote increased trade in value-added agricultural commodities produced in developing countries by assisting them in improving post-harvest technology and sanitary and phytosanitary measures.

5. Contribute generously to global emergency food security reserves and to Food for Livelihood-Security Programmes through the World Food Programme and other appropriate global, regional and national protective social security programmes, so that every child, woman and man has the requisite entitlements for a healthy and productive life.

I would suggest that the State of Food and Agriculture Report for 1995 could include in its special topic section a discussion on the scope and need for such a global convention for the elimination of hunger.

M. S. SWAMINATHAN

M. S. Swaminathan Research Foundation,
3rd Cross Street,
Taramani Institutional Area,
Madras 600 113, India.

A visionary and courageous scientist

An obituary of Jonas Salk

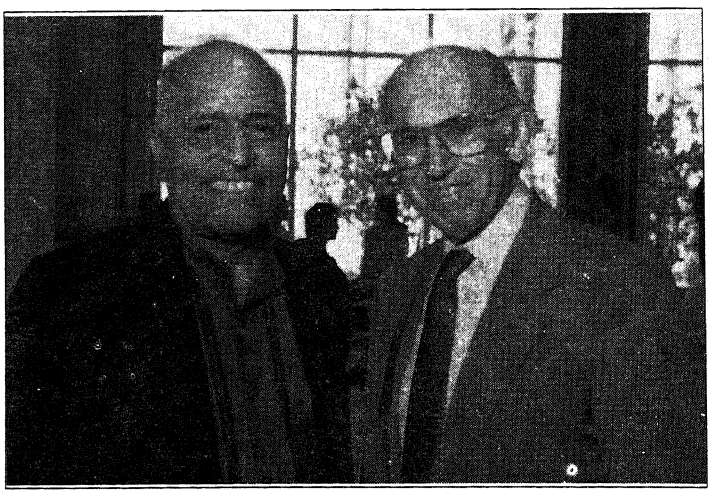
Jonas Edward Salk, a visionary and courageous scientist who invented the first vaccine to protect children (and adults) against the crippling disease poliomyelitis, passed away in La Jolla on 23 June at the ripe age of 80. His contributions are monumental and in countries continuously using the Salk vaccine the result was an initial drastic reduction in the incidence of the disease and, thereafter, total eradication of polio. He was also a builder of institutions and had to his credit the founding of the well-known Salk Institute for Biological Research.

His incursions in vaccinology began in 1942 at the University of Michigan, where he helped develop a commercial vaccine against the influenza virus. Several thousand soldiers in the US Army were immunized with this vaccine employing a potent immunopotentiator similar to Freund's incomplete adjuvant. A follow-up of these subjects 25 years later did not indicate any incidence of cancer or abnormality attributable to this immunization. Based on these observations he questioned the wisdom of not using such adjuvants for obtaining high and long-duration response with vaccines destined for human use. Before his death, he was deeply involved in making a vaccine against HIV in which he was intending to use this adjuvant.

In 1947, Jonas Salk moved to Pittsburgh in a faculty position as the Head of the Virus Laboratory. It is here that he started work on the polio vaccine, which received impetus from President Roosevelt, who had himself been a victim of polio. His researches were generously funded by public and private funds, prominent amongst which was a grant of $ 7.5 million from the March of Dimes campaign launched by the National Foundation for Infantile Paralysis. In an amazingly short period of 8 years, in a promotional climate free from obstructions, his vaccine passed from laboratory to human trials. The results of the large-scale trial announced in 1955 were impressive. The vaccine conferred 80–90% protection. Imbued with humanitarian concern, he did not patent the vaccine nor made money from it. It was licensed to six companies on easy terms for large-scale availability to the public. By 1960, the incidence of polio declined by 92% in the United States. President Eisenhower awarded him the Congressional Medal for Distinguished Civilian Achievement. It was taken up for immunization of children in other countries with equally impressive protection rates and fall in incidence of the disease. The French Govt. conferred on him the Legion d'Honneur for his work of international significance.

Then came a setback for no apparent fault of Salk. One of the batches of the vaccine manufactured by a company failed to kill the virus completely and the vaccine caused unforeseen deaths of about 300 recipients. The government imposed a ban on this vaccine in USA, even though it continued to be in use in Canada, Holland, Scandinavia and some other European countries. Meanwhile, a rival vaccine developed by Albert Sabin became available and was introduced in USA in 1961 based on trials conducted in USSR. The Sabin vaccine consisted of three strains of live *attenuated* polio viruses, in contrast to the Salk vaccine, which employed *killed* viruses. The Sabin vaccine had the ease of oral administration, whereas the Salk vaccine required to be injected. The counterpoints were that the Sabin vaccine demanded to be kept frozen to remain active, whereas the Salk vaccine did not require the 'cold chain'. The mode of action of the two vaccines was also somewhat different. The Salk vaccine, by systemic immunization, generated antibodies and cytotoxic cells, inactivating and killing the infecting polio virus. No live virus was released in the stools. The Sabin vaccine sought to induce local immunity in the gastrointestinal tract against the infectious virus. The live attenuated polio viruses present in the vaccine, however, colonized in the gut and were released to the environment in faeces. It was conceived that these viruses would be taken by others in the community and would hence create 'herd' immunity over and above the individuals immunized.

On face value, the Sabin vaccine was cheaper and appeared to be more suitable for general immunization programmes in children. However, trials conducted by Dr S. Bhallaya at the All India Institute of Medical Sciences, New Delhi, and by Dr Jacob John at the Christian Medical College, Vellore, showed the superiority of the Salk vaccine for inducing antibodies against the polio viruses as compared to three doses of the Sabin vaccine given orally. Similar results were reported from Algeria. The relative poor immunogenicity of the Sabin vaccine was not ascribable to the improperly frozen vaccine in the above-mentioned trials conducted by experienced virologists. It is surmised that five doses of the oral vaccine may be necessary for satisfactory sero-conversion of the children.



Jonas Salk (right) with the author at the 5th International Congress of Immunopharmacology at Tampa, Florida (1991).

This does pose practical difficulties of repeat visits and compliance in the expanded programmes of immunization. These difficulties are further compounded by the uncertainties of availability and maintenance of cold chain for this vaccine in far-flung rural areas. Thus, after due scrutiny of the data, the Indian Council of Medical Research recommended the use of the Salk vaccine as such or as adjunct to the Sabin vaccine. An opportunity to make the Salk vaccine in the country, along with two other cell-culture-derived vaccines against measles and rabies, came forth in 1989 with the French company Merieux agreeing to be a partner in a venture along with the Department of Biotechnology, Indian Petrochemicals, and 25% shares accessible to public for investment. Jonas Salk (and others including me) played an active role in motivating Merieux to come to India with a technology that will be useful for more than one vaccine of proven efficacy. An agreement was signed in New Delhi by the President of Merieux with counterpart officials in the august presence of President François Mitterand and Prime Minister Rajiv Gandhi. Had this agreement materialized, India would have been the largest producer of these vaccines in the world, meeting its own needs and exporting to other countries. Merieux had a further proposal of making a quadruple vaccine comprised of DPT (diphtheria, pertussis and tetanus) along with the Salk polio vaccine. Thus, complete immunization of the child against these four killer diseases could be accomplished with a common vaccine. A cherished goal of researchers and public health organizations is to have a vaccine with multiple components to prevent all childhood infections by a single intervention. Technologies are evolving to deliver multiple doses of the vaccine in biodegradable microspheres at a single contact point.

Jonas Salk was a pioneer vaccinologist. He was also a thinker and theorized on the implications of scientific findings on man and his place in society. He authored four books in this context: *Man Unfolding* (1972); *The Survival of the Wisest* (1973); *World Population and Human Values* (1981) and *Anatomy of Reality* (1983). He rallied readily for public causes. In spite of specialized focusing that any notable scientific work demands, his outlook was broad. He had a sensitive personality, and was at home amongst writers, artists and thinkers. He married in 1970 an equally lovable person, Françoise Gilot, an artist and a writer (*My Years with Picasso*) who was for many years a companion of and lived with Pablo Picasso. He is survived by three sons from his first wife.

Although Jonas Salk did not get the Nobel prize, he received several recognitions. In 1975, he delivered in New Delhi the Jawaharlal Nehru Award Oration for International Understanding.

In response to the havoc that AIDS was causing worldwide, he got energized in the last years of his life to make a vaccine against this 'clever' virus, which attacks and incapacitates the host immune status. With his demise, his ambition remains unfulfilled.

G. P. TALWAR

International Centre for Genetic Engineering and Biotechnology and The National Institute of Immunology, Aruna Asaf Ali Marg, New Delhi 110 067, India.

A doyen of Indian physicists

An obituary of S. D. Chatterjee

Professor Shyamadas Chatterjee, a doyen of Indian physicists, passed away in Calcutta on 27 May 1995 at the age of 86.

Born on 29 June 1909, he studied at Cuttack, Patna and then at Presidency College, Calcutta, obtaining his B Sc degree in 1930 and M Sc degree in Physics from Calcutta University in 1932. Among his teachers was none other than C. V. Raman, who on winning the Nobel Prize in 1930 graciously acknowledged the letter of congratulation from his young student. Chatterjee joined the Bose Institute, founded by the pioneer Jagadish Chandra Bose in 1938, and showed his experimental skills by constructing one of the first Wilson cloud chambers in India. For his research on cosmic rays he was awarded the D Sc degree of Calcutta University in 1945, his examiners being Blackett and Cockroft.

When the news of the discovery of nuclear fission reached India, he set up his own experiment with indigenous components to detect the emitted neutrons. He was thus the first one to detect the spontaneous fission of uranium in 1940. For the interpretation of his experimental results he was indebted to S. N. Bose. His brief communication to *Science and Culture* was, however, withdrawn as it had been sent without the permission of the then Director of the Bose Institute! His estimate of the half-life of uranium was a few orders of magnitude less than the value of Teller but was verified by Georgy Flerov from the Soviet Union within a few months. The actual calculation of the half-life from the detector counts was done by N. R. Sen of the Department of Applied Mathematics. In return Chatterjee repaired Sen's radio set, an indication of how scientific collaboration worked in those days! Chatterjee was awarded the Woodburn Medal of Calcutta University for best research in 1948 and the Elliott Medal of the Royal Asiatic Society in 1951.

In 1949 Chatterjee went to Chalk River, Canada, and also visited several laboratories in the USA, including the National Bureau of Standards, where he interacted with Tuve. He maintained contact with

leading scientists like Herzberg, Marton and Mott throughout his life.

Satyendra Nath Bose had moved back to Calcutta from Dacca in 1945 and recruited Chatterjee as Lecturer in the Department of Pure Physics in 1951. He was later promoted to Reader. He set up the first radiocarbon dating laboratory in India and detected fall-out across the country from atmospheric nuclear tests conducted in the early fifties.

Chatterjee was known for his brilliance in electronic instrumentation and built his own GM and proportional counters, picking up the coincidence counting technique from Walter Bothe in Germany. He worked as Fellow in The Davy–Faraday Laboratory of The Royal Institution, London, in 1958 when Lawrence Bragg was the Director. During one of his visits to Europe he is credited with the suggestion of the possibility of recoil-less emission of gamma rays from the nucleus, which led to the discovery of the Mössbauer effect. In recognition of this, he was invited to Munich as a Visiting Professor in 1964.

Chatterjee was invited to head the Physics Department of the newly constituted Jadavpur University in 1956 and built up the department with energy and devotion. His group discovered very high helium content from the emanations at the hot springs at Bakreshwar near Calcutta. For enriching the gas he set up his own laboratory which was later taken over by the Department of Atomic Energy. His research interests expanded into experimental solid-state physics, his students studying exo-electron emission from GM electrodes, surface states in semiconductors and metal–semiconductor contacts. He guided over 60 doctoral theses and was a kindly father-figure, ever ready to advise and support financially researchers in need.

On retirement from Jadavpur University in 1969, he joined his mentor S. N. Bose at the Indian Association for the Cultivation of Science as Senior Fellow and again set up a helium laboratory. He continued to guide research students, his encyclopedic knowledge of physics being always available to colleagues and students. A small laboratory was even set up in his own house and run by the S D Chatterjee Foundation.

He was elected Fellow of the American Physical Society and Fellow, IEEE and was one of the seniormost Fellows of the Indian National Science Academy. A bachelor with spiritual leanings, Chatterjee donated his house to the Asiatic Society Calcutta, one of the oldest institutions pursuing knowledge in all its aspects.

D. N. BOSE

Materials Science Centre,
Indian Institute of Technology,
Kharagpur 721 302, India.

A probe into the Calcutta earthquake of 1737

Pitta Govinda Rao

Some doubts have been expressed in the recent literature about the occurrence of an earthquake that occurred simultaneously with a cyclonic storm during September/October 1737 near Calcutta. In this article an attempt has been made to probe into the historical documents with a logical analysis. Our study reveals that an earthquake did occur undoubtedly near Calcutta in 1737. However, the death toll reported in the catalogue of significant earthquakes by the US Department of Commerce needs to be corrected. It was also found that the Encyclopedia Britannica has missed to make entry of the earthquake, in addition to making inadvertently a wrong entry of the date of occurrence of the cyclonic storm. This study has led to the conclusion that an earthquake did occur near Calcutta during the passage of a severe cyclonic storm on the Bengal coast on the night of 11 October 1737 as per the Gregorian equivalent calendar, or on 30 September 1737 as per the then-existing Julian calendar. The death toll due to the combined disaster could have been about 3000.

The United States Department of Commerce in their catalogue[1] of significant earthquakes for the period 2000 BC–1979 has made an entry of an earthquake occurrence near Calcutta on 11 October 1737. Recently, some scientific groups[2,3] have expressed doubts about the occurrence of this event. The suspicion was mainly due to the death toll reported in this document and the mismatch of the dates of occurrence in other sources of literature. Due to the coincidence of the earthquake event with a severe cyclonic storm that crossed the Bengal coast[4], it is felt necessary to investigate further to make the picture more transparent. In the process, the author has reviewed all the available literature related to the earthquake from the recent to the historical times[5]. An independent approach was adopted and reported briefly earlier[6]. In this article, an attempt has been made to give detailed information on the event. The author was fortunate to get all the necessary literature at the University of East Anglia, UK.

Main reports

The US Department of Commerce in their catalogue[1] of significant earthquakes, published in 1981, have included the occurrence of an earthquake on 11 October 1737 near Calcutta. The location was given as 22.5°N, 88.4°E and the number of deaths reported was 300,000. The damage caused has been rated as 'moderate', corresponding to a property loss of 1–5 million US dollars of 1979.

Encyclopedia Britannica of 1974 (ref. 7) mentioned only the occurrence of a cyclonic storm on 7 October 1737, causing a 12 m storm surge and killing about 300,000 people, and termed it as the most catastrophic cyclone.

Blanford, a meteorologist working in India, in his catalogue[8] of 1877 has included the occurrence of the cyclone storm on 7 October 1737 and described it as follows: 'This storm is ranked among the worst, if not the worst, tropical cyclone disasters in recorded history.' The death toll given by him was 300,000.

In 1848, Peddington[9] in his *Hornbook* stated clearly about the simultaneous occurrence of an earthquake and a cyclonic storm on 11 October 1737. He did not make any entry about the casualty figures.

Earlier, in 1837, Martin[10], in his *Medical Topography of Calcutta* writes:

> 'The earliest notice I can find of a storm and earthquake occurring together is – On the night between 11 and 12 October 1737, there happened a furious hurricane at the mouth of the Ganges which went 60 leagues up the river. There was at the same time a violet earthquake, which threw down a great many houses along the riverside in Golgota (i.e. Calcutta) alone, a port belonging to the English.... Dutch ships in the river, three were lost, with their men and Cargoes: 300,000 souls are said to have perished. The water rose 40 feet higher than usual in the Ganges.'

Gentleman's Magazine of June 1738 describes about the damage caused due to a furious hurricane in the Bay of Bengal and a violent earthquake on 30 September 1737. However, no death toll was included (a copy of the report from the magazine is given in Appendix 1).

Bengal Public Consultations, dated 15 October 1737, published in another source in 1906, states as follows:

> 'The late violent storm laid the whole black towns, quite throughout the honourable company's (The East India Company) bounds, insomuch that hardly twenty thatched houses were standing the next day, and the inhabitants lost the little they had and were rendered incapable of paying the rents ... what still adds to the calamity is that by the violent force of the wind the river overflew so much that a great quantity of rice was quite spoiled... and nearly 3000 inhabitants were killed, and a great number of large cattle besides goats and poultry destroyed.'

Discussion

It is evident from these reports that an earthquake did occur during September–October 1737. However, unfortunately, for reasons not known, reports earlier to 1974 do not provide the sources from where the information on the disaster was collected. If we analyse in the reverse order, i.e. from the historical times, we find that Martin[10] has nowhere indicated that he had collected his data from a journal published in 1738. He had added and exaggerated the information on death toll (from 3000 to 300,000) from an official source, *Bengal Public Consultations*, dated 15 October 1737. A close comparison of these sources clearly reveals that Peddington[9] reproduced the occurrence of the events from Martin's book[10]. In both Martin's and Peddington's reports, the date of occurrence of the storm and the earthquake was given as 11 October or the night between 11 and 12 October 1737. Blanford, without referring to any earlier reports, has reported the occurrence of a cyclonic storm on 7 October 1737, without any explanation regarding the date of the event[8]. Literature study reveals that he has taken the information from Peddington[9] and altered

CURRENT SCIENCE, VOL. 69, NO. 5, 10 SEPTEMBER 1995

THE

Gentleman's Magazine:

AND

Hiſtorical Chronicle.

VOLUME VIII.

For the YEAR M.DCC.XXXVIII.

CONTAINING,

I. Proceedings and Debates in the SENATE of *Lilliput*.

II. ESSAYS, Controverſial, Humourous and Satirical; Religious, Moral and Political: Collected in Part from the Publick Papers.

III. Diſſertations and Letters from Correſpondents.

IV. Select Pieces of POETRY.

V. A Chronicle of the moſt remarkable Tranſactions Foreign and Domeſtic.

VI. Births, Marriages, Deaths, and Promotions.

VII. The Prices of Stocks, Bill of Mortality, and Regiſter of Books.

PRODESSE & DELECTARE.



E PLURIBUS UNUM.

By *SYLVANUS URBAN*, Gent.

LONDON:

Printed by EDW. CAVE, *jun.* at St JOHN'S GATE.

Hiſtorical Chronicle, 1738.

JUNE.

ON *September* 30, laſt happened a furious Hurricane in the Bay of *Bengal*, attended with a very heavy Rain, which raiſed 15 Inches of Water in 6 Hours, and a violent Earthquake, which threw down abundance of Houſes; and as the Storm reached 60 Leagues up the River *Ganges*, it is computed that 20,000 Ships, Barks, Sloops, Boats, Canoes, *&c.* have been caſt away. A prodigious Quantity of Cattle of all Sorts, a great many Tygers, and ſeveral Rhinoceroſes were drowned; even a great many Caymans were ſtifled by the furious Agitation of the Waters, and an innumerable Quantity of Birds was beat down into the River by the Storm. Two *Engliſh* Ships of 500 Tons were thrown into a Village above 200 Fathom from the Bed of the River *Ganges*, broke to Pieces, and all the People drowned pellmell among the Inhabitants and Cattle. Barks of 60 Tons were blown two Leagues up into the Land over the Tops of high Trees. The Water roſe in all 40 Foot higher than uſual. The *Engliſh* Ships drove aſhore, and broke to Pieces, were the *Decker*, *Devonſhire*, and *Newcaſtle*; and the *Pelham* is miſſing. —— A *French* Ship was drove on Shore, and bulged; after the Wind and Waters abated they opened their Hatches, and took out ſeveral Bales of Merchandize, *&c.* but the Man who was in the Hold to ſling the Bales ſuddenly ceaſed working; nor by calling to him, could they get any Reply; on which they ſent down another, but heard nothing of him, which very much added to their Fear; ſo that for ſome Time no one would venture down: At length one more hardy than the reſt, went down, and became ſilent and unactive as the two former, to the Aſtoniſhment of all: They then agreed by Lights to look down into the Hold, which had a great Quantity of Water in it, and to their great Surprize, they ſaw a huge Alligator ſtaring as expecting more Prey: It had come in thro' a Hole in the Ship's Side, and 'twas with Difficulty they killed it; when they found the three Men in the Creature's Belly.

Wedneſday, 7.

On the 2d and 7th of this Month happened great Diſputes in the Court of Common Council of the City of *London*, concerning Propoſals for performing the Maſons Work of the intended Manſion-Houſe for the Lord Mayor; and on the Diviſions, one Party appeared in behalf of the Propoſal of Meſſ. *Townſhend*, *Horſenail*, and *Taylor*, three of their own Body, to do the ſaid Work for 17200 l. and the other Party for 250 l. cheaper. Some Perſons thinking it extraordinary that Gentlemen ſhould be for the higheſt Price, and alſo that it might carry a Reflection, as if thoſe propoſed to take the loweſt, were not able to do the Work ſo well as the others; inſerted in the News Papers the Oath of a Common-Councilman, diſtinguiſhing the Clauſe following in a different Character.

Good and true Counſel ye ſhall give in all things touching the Commonwealth of this City, after your Wit and Cunning: And that for Favour of any Perſon, ye ſhall maintain no ſingular Profit againſt the common Profit of this City. A Liſt alſo was handed about of 7 Aldermen, and 110 Commoners who divided for the higheſt Price, all, but about 20, of which Liſt are diſtinguiſhed to be of ſome of the Committees for adminiſtring the City Affairs; and, what is pretty remarkable, they are diſtinguiſhed, who voted on the firſt Diviſion for the loweſt Propoſal, *viz.* Sir *John Barnard*, (Lord Mayor) Sir *Robert Godſchall*, *Daniel Lambert*, and *Robert Weſtley*, Eſq; Aldermen. Dep. *Joſeph Ayliff*, Dep. *Tho. Sandford*, Dep. *Sam. Tatem*, Mr *Robert Henſhaw*, and Mr *Hen. Siſſon* Common Council-men.

The firſt Diviſion, *June* 2. was,

—— For the higheſt Price 80.

—— For the loweſt 63.

Tt The

some details and reported as independent information. He changed some quantitative information; for example, in the statement '... the storm reached 60 leagues up to the river Ganges ...' 60 leagues has been changed to 60 miles (a league is equivalent to about 3 miles).

Logical analysis reveals that *Encyclopedia Britannica*[7] contains the information given by Blanford, including the date of occurrence and the death toll[8]. Since Blanford's interest is only for the weather-related phenomena, he has excluded the earthquake occurrence. The report of the US Department of Commerce also mentions the death toll as 300,000 due to an earthquake on 11 October 1737. This document quotes several references[12–17] as sources of information. After a thorough study, the author is inclined to conclude that the source of information for these reports is also the report of Martin[10].

The next question which arises is about the date of occurrence. It is evident that after a century of occurrence, Martin reported[10] the date of the events of cyclone and earthquake as the night of 11 October 1737. It is important to note that he has taken care of the change of the calendar by the Britishers. The Gregorian calendar was not in use in British Empire in 1737 and the date 30 September is according the Julian calendar then in vogue. It matches very well with the date 11 October 1737, which would be its Gregorian equivalent[4], but not the 7 October as mentioned by Blanford. The present author, while looking at the historical source[5], *Gentleman's Magazine*, was astonished to see the total absence of any estimate of lives lost in the combined disaster (Appendix 1). This led him to search further and find the official figure of death toll in the combined tragedy as 3000 in another source[11]. In this, a letter written more than a year after the disaster, on 29 January 1739, describes:

> '... the storm which levelled most of the walls in the town, shattered and threw down many of the buildings and blew up the bridges, the tide some days after broke in upon and carried away some of the wharfs, ships and stairs, the places most damnified are the peers on the factory wharf, wharf and ships at Soota Loota. Walls around the burying place and powder magazine and factory points, Church steeple was overthrown.'

The above description enriches the knowledge for supporting the occurrence of an earthquake along with the storm. Further, in an economic study of Bengal, Bhattacharya[18] mentions that only after the great disaster of 1737 was widespread masonry construction considered desirable. These observations give a fairly good qualitative idea that destruction to buildings might have occurred partly due to ground motion. Bilham[2], while concluding that 'the listing of this event in the catalogues of disastrous earthquakes is indefensible', writes that 'the possibility that an earthquake-induced *tsunami* caused the flood damage cannot be completely excluded'.

Population estimates range from 10,000 to 12,000 between 1705 and 1720 (ref. 19). It has been documented[20] based on the report by Ray[21] that the population of Calcutta was 120,000 in 1757. It was further mentioned that 'the early nineteenth century may not have held for the 200,000 or so who were living in Calcutta, in an environment largely of British creation'[20]. From these facts it is clear that the number of people dead in the disaster *cannot* be 300,000 as reported by the *Encyclopedia Britannica* and by the US Department of Commerce. The author, therefore, feels that the official figure of 3000 might have been exaggerated to 300,000 by Martin[10].

Conclusion

After spending sufficient time in reviewing the literature from the historical past to the present, the author feels confident in stating that an earthquake did occur in the vicinity of Calcutta on the night of 11 October 1737 as per the Gregorian equivalent calendar, or on 30 September 1737 as per the then-existing Julian calendar. Exaggerating the death toll figures is a human psychology, even today, for drawing the attention of the targeted groups. It is also evident that the earthquake occurrence was during the period of landfall of a cyclonic storm on the Bengal coast. The concerned authorities of the *Encyclopedia Britannica* and the US Department of Commerce may consider issuing a modification to their reports, after another independent study, if felt necessary.

1. Ganse, A. R. and Nelson, J. B., *Catalog of Significant Earthquakes 2000 BC–1979*, US Department of Commerce, Boulder, CO, 1981, p. 17.
2. Bilham, R., *Bull. Seismol. Soc. Am.*, 1994, accepted.
3. Hecht, J., *New Scientist*, 1994, **142** (1921) 7.
4. Sensarma, A. K., *Weather*, 1994, **49**, 90–96.
5. *The Gentleman's Magazine and Historical Chronicle* (ed. Sylvanus, U.), Cave, E. Jun at St John's Gate, London, 1738, **VIII**, June, p. 331.
6. Rao, P. G., *New Scientist*, 1994, June, p. 51.
7. *The New Encyclopaedia Britannica*, Chicago, USA, 1974, vol. 9, p. 58.
8. Blanford, H. F., *J. Asiatic Soc. Bengal*, 1877, **46** (Pt II), 328.
9. Peddington, H., *Sailors Hornbook for the Law of Storms*, Smith Elder, London, 1848, p. 175.
10. Martin, J. R., *Notes on the Medical Topography of Calcutta*, Guttmann, Calcutta, 1837, pp. 7–8.
11. Wilson, C. R., *Indian Records Series*, John Murray, London, 1906, vols 1, 2.
12. Bath, M., *Introduction to Seismology*, Halsted Press, Basel, 1978.
13. Lomnitz, C., *Development in Geotectonics*, Elsevier, 1974, vol. 5.
14. Myers, H. M. and Von Hake, C. A., *Earthquake Data File Summary*, National Geophysical and Solar-Terrestrial Data Centre, Boulder, CO, 1976.
15. Münchener Ruckvericherungs-Gessellschaft (MRG), *World Map of Natural Hazards*, Munich, 1978.
16. Milne, J., *Catalogue of Destructive Earthquakes*, Report of the 18th Meeting of the British Association for the Advancement of Science, Portsmouth, London, 1912, pp. 649–740.
17. Singh, K., *Indian J. Meteorol. Geophys.*, 1966, **17**, 4.
18. Bhattacharya, S., *The East India Company and the Economy of Bengal from 1704 to 1740*, Luzac, London, 1954.
19. Hamilton Alexander, *A New Account of the East Indies*, 2, Edinburgh, 1927.
20. Marshall, P. J., *The New Cambridge History of India*, Cambridge University Press, pp. 159, 172.
21. Ray, A. K., *Census of India 1901*, Calcutta, 1902, **VII**, 59–62.

ACKNOWLEDGEMENTS. I thank Prof. N. N. Ambraseys of Imperial College of Science and Technology, London, and Dr Paul W. Burton of the University of East Anglia, UK, for help in literature search.

Pitta Govinda Rao is in the School of Environmental Sciences, University of East Anglia, Norwich, NR4 7TJ, UK. Permanent affiliation: ESS Division, Department of Science and Technology, New Delhi 110 016, India.

MINISTRY OF SCIENCE AND TECHNOLOGY
DEPARTMENT OF SCIENCE AND TECHNOLOGY

JSPS RONPAKU (DISSERTATION Ph D) FELLOWSHIPS: 1996–97

Department of Science and Technology (DST) invites applications from promising and qualified young researchers for JSPS RONPAKU (Dissertation Ph D) Fellowships: 1996–97, offered by the Japan Society for the Promotion of Science (JSPS) within the framework of India–Japan Cooperative Science Programme (IJCSP). Four Fellowships are available.

Description of the Ronpaku programme

RONPAKU programme provides support to researchers who wish to obtain their Ph D. degrees from Japanese universities through the submission of dissertation, without going through the course work.

Grantees under this RONPAKU programme (RONPAKU Fellows) are required to conduct dissertation research under the joint supervision of a Japanese Advisor and an Indian Advisor. For this purpose, during the maximum period of five years, the RONPAKU Fellow will visit Japan once a year to conduct research at a Japanese university under the supervision of the Japanese Advisor, who will visit the RONPAKU fellow's country, when necessary, to supervise the research together with the fellow's home advisor.

Fields

All fields of Science & Technology

Eligibility

Applicant should

1. be a citizen of India;
2. have already achieved substantial research results and, through dissertation research under the programme, be deemed likely to obtain a Ph D degree successfully;
3. be holding a regular position as a researcher in a recognized scientific/technical institution in India;
4. be able to forward the final application through proper channel of the Parent Institution.

Japanese Advisor

The Japanese Advisor must be either a professor, associate professor or lecturer who teaches in the doctoral course of a graduate school of either a national, public or private university in Japan.

Supervision of research

During the overall tenure of 5 years, the RONPAKU fellow will visit Japan once a year for a maximum period of 90 days per visit, to conduct research under the supervision of the Japanese Advisor.

The Japanese Advisor may visit the fellow's country, not more than once a year (for a period of 7–30 days per visit) to supervise the research in accordance with the plan approved.

Terms of awards

JSPS provides the following financial support:

To the RONPAKU fellow

1. Round-trip economy class air ticket between India and Japan.
2. Maintenance allowance as per JSPS regulations.
3. In Japan, research trip allowance of up to Yen 100,000 per visit to Japan once a year.
4. Sickness and accident insurance fee during the stay in Japan, as per JSPS regulations.
5. Expenses related to dissertation submission and publication in the final year of the fellowship tenure, as per JSPS regulations.

To the Japanese Advisor [when visiting the fellow's country]

1. Round-trip economy class international airticket between Japan and India.
2. Maintenance allowance for stay in India.
3. Supervising expenses (when receiving a RONPAKU fellow).
4. Research allowance for supervising the RONPAKU fellow.
5. Supervising expenses.

Application procedures

The Application Forms can be obtained from the Programme Office, India–Japan Co-operative Science Programme (IJCSP), International Division, Department of Science & Technology, Technology Bhavan, New Mehrauli Road, New Delhi 111 016 by sending a request along with a self-addressed envelope (12 cm × 27 cm) entitled "JSPS RONPAKU Fellowship 96–97."

The filled application forms should be forwarded by the candidate through proper channel latest by 15 October 1995. INCOMPLETE APPLICATIONS WILL BE SUMMARILY REJECTED. Applicants will be notified of the results by the end of March 1996.

CURRENT SCIENCE

A fortnightly journal of research

Editors: S. Ramaseshan and P. Balaram

Editorial Board

S. Arunachalam, Central Electrochemical Research Institute, Karaikudi 623 006
Asis Datta, School of Life Sciences, Jawaharlal Nehru University, New Delhi 110 067
D. Balasubramanian, Centre for Cellular and Molecular Biology, Hyderabad 500 007
J. Chandrasekhar, Department of Organic Chemistry, Indian Institute of Science, Bangalore 560 012
R. Gadagkar, Centre for Ecological Sciences, Indian Institute of Science, Bangalore 560 012
K. N. Ganeshaiah, Department of Plant Genetics and Breeding, University of Agricultural Sciences, Bangalore 560 065
R. V. Hosur, Tata Institute of Fundamental Research, Bombay 400 005
N. V. Joshi, Centre for Ecological Sciences, Indian Institute of Science, Bangalore 560 012
Jyotirmoy Das, Indian Institute of Chemical Biology, Calcutta 700 032
C. C. Kartha, Division of Cellular and Molecular Cardiology, Sree Chitra Tirunal Institute for Medical Sciences and Technology, Thiruvananthapuram 695 011
S. S. Krishnamurthy, Inorganic and Physical Chemistry Department, Indian Institute of Science, Bangalore 560 012
S. C. Lakhotia, Department of Zoology, Banaras Hindu University, Varanasi 221 005
K. Muralidhar, Department of Zoology, Delhi University, Delhi 110 007
R. Narasimhan, CMC Limited, Bangalore 560 001
K. Neelakantan, ANURAG (Advanced Numerical Research and Analysis Group), Hyderabad 500 258
R. Nityananda, Raman Research Institute, Bangalore 560 080
T. Padmanabhan, Inter-University Centre for Astronomy and Astrophysics, Pune 411 007
B. R. Pai, Propulsion Division, National Aerospace Laboratories, Bangalore 560 017
T. J. Pandian, School of Biological Sciences, Madurai Kamaraj University, Madurai 625 021
Y. S. Rajan, Department of Science and Technology, New Delhi 110 016
M. Ramakrishnan, Geological Survey of India, Calcutta 700 016
A. V. Ramani, T. T. K. Pharma Ltd., Bangalore 560 025
S. Ramaswamy, Department of Physics, Indian Institute of Science, Bangalore 560 012
A. K. Raychaudhuri, Department of Physics, Indian Institute of Science, Bangalore 560 012
S. R. Shetye, Physical Oceanography Division, National Institute of Oceanography, Dona Paula, Goa 403 004
V. Siddhartha, 51 Bharati Nagar, New Delhi 110 003
V. Srinivas, School of Mathematics, Tata Institute of Fundamental Research, Bombay 400 005
G. Srinivasan, Raman Research Institute, Bangalore 560 080
R. Uma Shaanker, Department of Crop Physiology, University of Agricultural Sciences, Bangalore 560 065
K. S. Valdiya, Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore 560 064
G. Venkataraman, Department of Physics, Sri Sathya Sai Institute of Higher Learning, Prasanthi Nilayam 515 134
M. Vidyasagar, Centre for Artificial Intelligence and Robotics, Bangalore 560 001
K. VijayRaghavan, TIFR Centre, Indian Institute of Science Campus, Bangalore 560 012

Current Science, founded in 1932, is published by the Current Science Association in collaboration with the Indian Academy of Sciences. The journal is also intended as a medium for communication and discussion of important issues that concern science and scientific activity. All articles published in *Current Science*, especially editorials, opinions and commentaries, letters and book reviews, are deemed to reflect the individual views of the authors and not the official points of view, either of the Current Science Association or of the Indian Academy of Sciences.

Working Committee, Current Science Association

P. Balaram, Indian Institute of Science, Bangalore
R. Chidambaram, Department of Atomic Energy, Bombay
V. K. Gaur, CSIR Centre for Mathematical Modelling and Computer Simulation, Bangalore
K. Kasturirangan, ISRO Satellite Centre, Bangalore
P. K. Kaw, Institute for Plasma Research, Gandhinagar
N. Kumar, Raman Research Institute, Bangalore (*Vice-President*)
N. V. Madhusudana, Raman Research Institute, Bangalore (*Secretary*)
A. Mani, Field Research Unit, Indian Institute of Tropical Meteorology, Bangalore (*Past-President*)
R. A. Mashelkar, Council of Scientific and Industrial Research, New Delhi
R. Narasimha, Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore
G. Padmanaban, Indian Institute of Science, Bangalore (*Vice-President*)
R. Ramachandran, The Institute of Mathematical Sciences, Madras
P. Rama Rao, Defence Research and Development Organization, New Delhi
S. Ramaseshan, Raman Research Institute, Bangalore (*Treasurer*)
M. S. Valiathan, Manipal Academy of Higher Education, Manipal (*President*)

Editorial Office

Editorial staff

Chandrika Ramesh, G. Madhavan, T. D. Mahabaleswara, C. S. Ravi Kumar, M. S. Venugopal

Circulation and Accounts

Peter Jayaraj, Ranjini Mohan, B. Sethumani, Shanthi Bhaskar, B. K. Shivaramiah

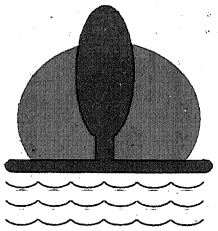
Annual subscription 1995

	Institutional	Personal
India	Rs 350	Rs 100
All other countries (AIR MAIL)	US $175	US $50

Single copies other than special issues: Rs 35/US $15

Editorial Office: CURRENT SCIENCE
C. V. Raman Avenue
P.B. 8001
Bangalore 560 080, India
Telephone: 91-80-3342310
Telex 845-2178 ACAD IN
Fax: 91-80-3346094

Copyright © 1995, Indian Academy of Sciences



IPS-11

July 28–Aug 02, 1996
Bangalore, India

Eleventh International Conference on Photochemical Conversion and Storage of Solar Energy (IPS-11)

Scientific Programme

The IPS conferences are held every two years with the aim of outlining the latest developments in the field of photochemical conversion and storage of solar energy. The eleventh conference of this kind will cover the following areas:

* Photoinduced electron transfer processes
* Photocatalysis in heterogeneous and homogeneous systems
* Photoelectrochemistry
* Photosynthesis in biological and biomimetic systems
* Quantum chemical studies of molecules on surfaces under the influence of light
* Thermodynamic aspects of quantum solar energy conversion
* Environmental applications of photocatalytic systems
* Materials for photoenergy conversion and photochemical engineering

These topics will be discussed in plenary and at invited lectures, workshops and poster sessions. Contributions from participants are encouraged. A book of abstracts will be distributed at the conference.

Location

The Eleventh International Conference on Photochemical Conversion and Storage of Solar Energy (IPS-11) will be held in the National Science Seminar Complex, Bangalore, India.

International Organizing Committee

E. Amouyal (France): C.A. Bignozzi (Italy): M. Calvin (USA); G. Calzaferri (Switzerland): Y. Cao (China); M. A. Fox (USA); M. Fujihira (Japan); D. Gust (USA); G. Hodes (Israel); V. Krishnan (India); J.-M. Lehn (France); N. S. Lewis (USA); S. E. Lindquist (Sweden); A. Nozik (USA); V. N. Parmon (Russia); E. Pelizzetti (Italy); Sir G. Porter (UK); N. Serpone (Canada); H. Tributsch (Germany); F. Wilkinson (UK); H. Yoneyama (Japan); F. Willig (Germany); K. I. Zamaraev (Russia)

Local Organizers

S. Umapathy and P. K. Das, Department of Inorganic and Physical Chemistry, Indian Institute of Science, Bangalore.

Second Circular

The second circular containing more details will be sent by October 1995. The Organizing Committee would appreciate your supplying this information to others who may be interested in attending the IPS-11.

Correspondence

The Secretariat IPS-11
Prof. V. Krishnan
Department of Inorganic and Physical Chemistry
Indian Institute of Science
Bangalore 560 012, India
Tel: 91-80-3092296
Fax: 91-80-3341683.

Founded: 1932

ISSN: 0011-3891

CURRENT SCIENCE

Volume 69 Number 6

25 September 1995

483 **In this issue**

NEWS

485 **ISI's electronic library**
Subbiah Arunachalam

486 **Total solar eclipse – October 1995**
J. C. Bhattacharyya

489 **10th Himalaya–Karakoram–Tibet Workshop, Switzerland, 1995 – Conference report**
Rasoul B. Sorkhabi

CORRESPONDENCE

491 **A good beginning**
J. S. Virdi

RESEARCH NEWS

491 **Does low serum cholesterol increase suicide and aggressive behaviour?**
Amar Bavle

492 **Bose–Einstein condensation in a dilute atomic vapour**
N. Kumar

SCIENTIFIC CORRESPONDENCE

493 **'Swiss' strain of rats?**
N. Giridharan

494 **Regeneration of *Aegle marmelos* (L.) Corr. plants *in vitro* from callus cultures of embryonic tissues**
R. Islam, M. Hossain, M. R. Karim and O. I. Joarder

495 ***In vitro* axillary bud break and multiple shoot production in *Acacia auriculiformis* by tissue culture technique**
P. Chandrasekhara Reddy, Veeranagouda Patil, T. G. Prasad, K. Padma and M. Udayakumar

496 **Effect of feotal exposure to low-dose X-rays on the postnatal growth of mouse**
K. S. Bisht and P. Uma Devi

GENERAL ARTICLES

499 **The quest for quantum gravity**
Gary K. Au

518 **Faint star counts and the Milky Way structure**
Devendra Ojha

REVIEW ARTICLE

521 **A unique and remarkable binary pulsar**
Simon Johnston

RESEARCH COMMUNICATIONS

526 **A high-precision technique using X-ray reflectivity for the measurement of surface and interface roughness**
P. V. Satyam, D. Bahr, S. K. Ghose, G. Kuri, B. Sundaravel, B. Rout and B. N. Dev

529 **Single-ion activities by a solid ion transmitter bridge and a reference electrode without liquid junction**
S. Parthasarathy and K. Ramya

534 **Fall of liquid drops through a viscious medium: A fresh expression for drag force by the method of dimensions**
G. J. Srinivasan and P. Satyanarayana

538 **Antibody response to a potentially protective antigen in human filariasis**
Madhusmita Bal and Manoj K. Das

541 **Visual demonstration of *in vivo* acid phosphatase activity of VA mycorrhizal fungi**
J. C. Tarafdar

543 ***In vitro* flowering and pod formation in cauliflower (*Brassica oleracea* var. *botrytis*)**
Vandana A. Kumar, Atul Kumar and Jitendra Kumar

545 **Testosterone biosynthesis in triploid sterile male tilapia**
R. Koteeswaran, S. G. Sheela and T. J. Pandian

BOOK REVIEWS

548 **Ethnobotany and the Search for New Drugs,** *reviewed by* Subbiah Arunachalam ■ **549 Environmental Health Perspective: Supplements: Molecular Mechanisms of Metal Toxicity and Carcinogenicity,** *reviewed by* Prasanta Kumar Ray ■ **550 Annual Review of Biochemistry 1994,** *reviewed by* K. Muniyappa

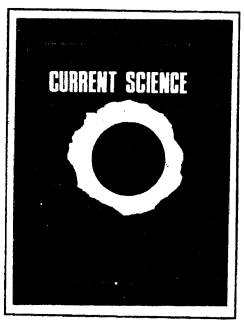
PERSONAL NEWS

552 **Joseph Needham (1900–1995),** *an obituary by* A. Rahman

554 **Subrahmanyan Chandrasekhar,** *an obituary by* R. Nityananda

HISTORICAL COMMENTARY AND NOTES

557 **On Ramanujan's bust**
S. Chandrasekhar



COVER. Solar corona of 16 February 1980, photographed with a 6-m focal length camera at the Hosur Camp, 40 km South of Hubli. Exposure: 30 s; mean epoch: 15 h, 41 m, 10 s IST. (Photo: IIA, Bangalore). See page 486.

Indexed in CURRENT CONTENTS/GEOBASE/CHEMICAL ABSTRACTS

In this issue

Quest for quantum gravity

This issue carries an article by Gary K. Au (**page 499**) on the 'Quest for Quantum Gravity'. It is an extended exposition of some current efforts towards the Holy Grail of theoretical physics. The article makes a real effort to aim at the reader outside the field, and is also unusual in avoiding the impersonal style of the textbook and the research journal, by using detailed extracts from interviews with three of the many people who have spent time on this quest – Chris Isham of Imperial College, London, Abhay Ashtekar of Pennylvania State University, and Edward Witten of the Institute of Advanced Studies, Princeton. We therefore have an interesting contrast in emphases, viewpoints and style. The article does not portray the field as a finished product but is a report from the frontline, so to speak. It will be interesting to see how much of what is described will seem outdated ten or even five years from now, and how much will survive.

R. Nityananda

Fall of liquid drops

The phenomenon of a liquid drop falling under gravity in another immiscible liquid medium is, in fact, a complex one. In general, the drop may deform considerably or even break up, the motion may be unsteady or oscillatory and the wake may even be turbulent. For sufficiently small drops, however, these pathologies are absent and the motion is steady with only slight deformation of the drops. A natural question to ask then is: What is the nature of the drag force on a drop under these circumstances?

In this issue (**page 534**) G. J. Srinivasan and P. Satyanarayana present the results of a detailed experimental study aimed at determining the drag force on a falling drop. They show that although, in general, many more dimensionless groups are involved, for the range of Reynolds numbers, Morton numbers and Eotvos numbers considered, only three are really relevant. Moreover, using data obtained from 21 liquid systems, they show that the relationship between these three yields an explicit expression for the drag force, a quantity of considerable practical interest. One should note that this expression holds even when the drops deform slightly provided the equivalent diameter is used. The authors point out that their results provide a means for measuring the density of liquids available only in microquantities when other conventional methods may fail.

P. N. Shankar

Measuring single-ion activities

Solutions of electrolyte were considered non-ideal, since the change in chemical potential of these solutions was not found to be equal to $RT \ln m$. To account for this, Lewis proposed an empirical correction factor, γ, called the activity coefficient and considered $m\gamma$ to be the effective concentration. The reason for the non-ideal behaviour of electrolyte solution was explained as being due to electrical interaction between ions by the Debye–Huckel limiting law. Thus, the concept of activity coefficient of ions gained a rational basis. However, the activity of a single ionic species remained a theoretical idea since it could not be measured experimentally and the measurable mean ionic activity was used in theoretical calculations.

Activity coefficient is conveniently measured by the EMF of galvanic cells with or without transference. The former involves a junction between two solutions and hence a liquid junction potential (LJP). The latter also includes LJP due to the contact between the reference electrode solution and the electrolyte solution and the electrolyte of the half cell under test. Evaluation of LJP needs a knowledge of single-ion activity (SIA). Thus, one faces a vicious circle when an attempt is made to measure SIA.

In this issue (**page 529**) S. Parthasarathy and K. Ramya propose an interesting solution to overcome this problem. They eliminate liquid junction by using a bridge of solid silver chloride, which is a solid-state ionic conductor. Using this idea, they have measured SIA of Zn^{2+} and Cl^- in $ZnCl_2$ solutions and the results are discussed in this paper. An interesting observation from the results is that the activity is seen to be equal to $Z_i^2 mr_i$ and not mr_i, lending support to the view of Parthasarathy and Srinivasan published in *Current Science* (1994, **66**, 56–59).

S. R. Rajagopalan

Patrons

Dr. S. K. Joshi
President, INSA

Dr. R. A. Mashelkar
Director General, CSIR

Dr. G. V. Satyabati
Director General, ICMR

Dr. C. R. Bhatia
Secretary, DBT

Dr. V. S. Ramamurthy
Secretary, DST

Advisory Committee

President
Prof. V. Ramalingaswami

Members
Prof. B. K. Bachhawat
Prof. L. S. N. Prasad
Prof. G. Padmanabhan
Prof. B. N. Dhawan
Prof. Indira Nath
Dr. Sandip K. Basu
Dr. Manju Sharma
Dr. Asish Mukherjee
Dr. V. P. Sharma
Prof. R. C. Mahajan
Prof. N. K. Ganguly
Dr. Lalit Kant
Dr. V. P. Kamboj
Prof. P. N. Tandon

Organising Committee

President
Prof. A. B. Chowdhury

Executive President
Prof. A. N. Bhaduri

Vice-Presidents
Prof. A. K. Hati
Prof. P. K. Ray
Dr. Santanu Kar
Dr. J. Das
Dr. D. Basu

Jt. Secretaries
Dr. M. K. Basu
Dr. H. K. Majumder

Treasurer
Dr. Dwijen Sarkar

Members
Prof. Asoke Ghosh
Prof. V. K. Vinayak
Prof. A. Nandy
Prof. A. Mukherjee
Prof. A. Bhattacharyya
Dr. Tuli Biswas
Dr. Mitali Chatterjee
Dr. Syamal Roy
Dr. Kawna Chatterjee
Shri Arun Banerjee
Smt. Debasree Mukherjee



SYMPOSIUM
on
U. N. BRAHMACHARI & PERSPECTIVE OF KALA-AZAR RESEARCH

Indian Institute of Chemical Biology
Calcutta 700 032

Date: 3–4 November 1995

TOPICS OF SCIENTIFIC SESSIONS

1. History of Kala-azar research and contributions of Professor Brahmachari
2. Clinical profile of the disease including PKDL
3. Epidemiology and transmission dynamics of the disease
4. Chemotherapeutic approach
5. Vector biology and control
6. Immunobiology of host–parasite interaction
7. Biochemistry and molecular biology of Leishmanial parasites
8. Pathophysiology of Leishmaniasis
9. Diagnosis
10. Control

Address for correspondence:

Dr. M. K. Basu/Dr. H. K. Majumder
Indian Institute of Chemical Biology
4, Raja S. C. Mullick Road, Calcutta 700 032
Phone: 473-6793/0492/5197, Fax: 473-5197/0284

CURRENT SCIENCE

Volume 69 Number 6 25 September 1995

NEWS

ISI's electronic library

First announced in August 1994, the Institute for Scientific Information's Electronic Library Project[1] has made considerable progress in its first year.

The project in which IBM and Lotus collaborate with ISI, aims to deliver complete information to customers' desktops. The prototype consists of *Current Contents/Life Sciences* edition, providing table of contents, bibliographic data, and abstracts for 1350 journals, with the addition of full images of those articles for which publishers have agreed to allow electronic storage and delivery.

All aspects of the prototype are now in place. IBM offers technical advice and components for storage, retrieval and delivery of information in a distributed environment which will seamlessly plug into various existing library WAN/LAN topologies. Lotus Notes Desktop software will be used by pilot sites at eight locations – six in North America and two in the UK – which have agreed to test the prototype.

ISI chose Lotus Notes as it is the leading client/server platform for developing and deploying groupware applications, and it has the capacity to deal with large amounts of full-imaged data and to track usage statistics. Besides, it is easily integrated into ISI's prototype, and provides the level of security needed for the content providers, viz. the journal publishers.

ISI has invited over 360 publishers, representing the 1350 journals indexed in *Current Contents/Life Sciences*, to participate in this venture. Two Indian journals, *Journal of Biosciences* and *Journal of Genetics*, both published by the Indian Academy of Sciences, are among the journals that have agreed to make full text images available to clients of ISI's electronic library. The eight pilot sites that would test the prototype are Brookhaven National Laboratory, Lehigh University, the New York Public Library, Purdue University, SmithKline Beecham Pharmaceutical R&D, Thomas Jefferson University, University College in London, and Glaxo Research and Development Ltd.

The test phase of the prototype has begun and will last for eighteen months. The test phase aims to shed light on practical applications such as data access, retrieval and usage. Input from library communities and publishers will help clarify issues such as copyright, intellectual property, pricing and usage patterns, that are key to the commercial success of the project.

The cornerstone of the project is the construction of an economic model that will facilitate pricing that will be acceptable to the information providers and the users. The project will measure the economics involved, weigh the need for security vs access, and will measure and assess usability. ISI's prototype electronic library can test:

1. Technical systems required to support electronic journal delivery (data access, storage and delivery)
2. Internal systems required to facilitate implementation of the electronic library – including billing, accounting and business management reporting
3. Economic models that will meet the diverse needs of the publishing and user communities
4. Behaviour of users to determine how the electronic journal may change the traditional information purchasing and usage patterns.

The economic model is based on several assumptions

1. The *Current Contents* database will be a site license.
2. Pricing for the electronic journals will be subscription-based and will be set by the publisher.
3. Royalty payments for documents ordered through the system will be set by the publisher. ISI will set a processing fee.

The prototype will provide for seamless network integration through a network-based client/server system that can be added to and coexist with the customer's current local area network. The ISI server at the pilot site will be capable of supporting heterogeneous clients, e.g. Windows, Macintosh, OS/2, etc. For reasons of security, data transmission will be via dedicated ISDN lines and not through Internet. The network uses the standard TCP/IP both for communication between Central Server and Customer Server and for traffic between Customer Server and Desktop.

In a way, the seeds of the idea of an electronic library were sown more than 50 years ago, when in 1954, Gene Garfield wrote as the first Grolier Society Fellow at Columbia University that 'the ideal library must be able to supply information instantaneously [and] satisfy a wide variety of information requirements for a population that will be highly intellectual and scientifically trained'. The founder of ISI then used the term 'library system for the year 2045!' Now it appears, thanks to developments in information technology and the growing demand for accurate information to meet specific demands, that his dream will come true 50 years earlier than anticipated by him and others.

1. Arunachalam, S., *Curr. Sci.*, 1994, 67, 977–978.

Subbiah Arunachalam, Central Electrochemical Research Institute, Karaikudi.

Total solar eclipse – October 1995

On October 24 this year, the track of totality of a solar eclipse will cross a 40–55 km wide strip of land in northern India. Total eclipse may be seen from any location within this strip, which will stretch from Rajasthan in the west to West Bengal in the east. The beginning of the eclipse will be over Iran at sunrise, and will end over the Pacific Ocean at sunset. In India, the total phase will be visible between 8.30 and 8.50 in the morning as the shadow will sweep over the country; the duration of totality on the centre line will range from 45 s over Rajasthan to about 80 s in West Bengal. Full details of the visibility of the eclipse have been calculated and given in the NASA Reference Publication No. 1344 (ref. 1). Weather conditions over the path are expected to be favourable[2].

One point about total solar eclipses often missed is that the view of the totally eclipsed sun cannot be obtained from outside the track of totality; what one can see is only a partially eclipsed sun. A few accompanying features such as considerable diminution of daylight, perhaps even a view of a few bright stars and planets may be seen, but most of the spectacles, the chromospheric flash, the glorious view of the corona, etc., will be missed. But these views are so rare that many who had stayed outside the track do not even realise what they have missed. I have come across many persons who claim to have witnessed a total eclipse, but on scrutiny it has come out that they have witnessed only a partial phase of the total eclipse. The records in such cases reveal that the track of totality of that particular eclipse did not pass through the city where they were located.

The chances of an eclipse track passing over a particular spot are very rare; statistical calculations indicate that on an average the chances are roughly once in 350 years. We may check the rate by comparing the visibility chances at a few particular spots. This year the track will pass slightly south of the city of Calcutta; the last time such an event was seen from this locality was on 10 May 1575 during the reign of Akbar, the Mogul. The last time an eclipse track passed within 50 km of Bangalore was on 5 July 1293; the city was not even founded. The next time it will be on 5 July 2168, a really long wait indeed!

But as in all statistical figures, certain extremes are unavoidable. In recent history, two eclipse tracks crossed over Newfoundland in the eastern coast of Canada within two years and four months, i.e. on 7 March 1970 and 10 July 1972. Still more strange is a spot in Siberia, where three eclipse tracks passed over within eleven years, viz. 1941, 1945 and 1952.

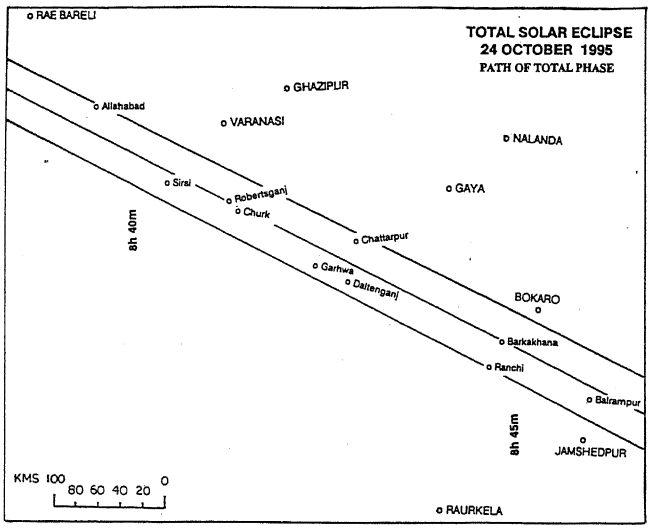
As far as the subcontinent of India is concerned, calculations reveal that during the twentieth century, eclipse tracks should pass over part of it just five times. Three events are already over; two, including the one this year, are yet to happen. The first was on 21 August 1914, when the track ended over the Rann of Kutch at sunset. The second, a similar one on 30 June 1954 had the path of totality slightly north, which again ended at sunset over the Rajasthan desert. The third one was the best among the five; its duration was about 3 min and the totality path crossed over Karnataka, Andhra Pradesh, and Orissa, and again over a short stretch over Mizoram. It happened in the afternoon with the sun still high in the sky and provided a unique opportunity of viewing the beauty of totally eclipsed sun to millions on its track. It is beside the point that countless numbers deprived themselves of this chance because of blind superstition.

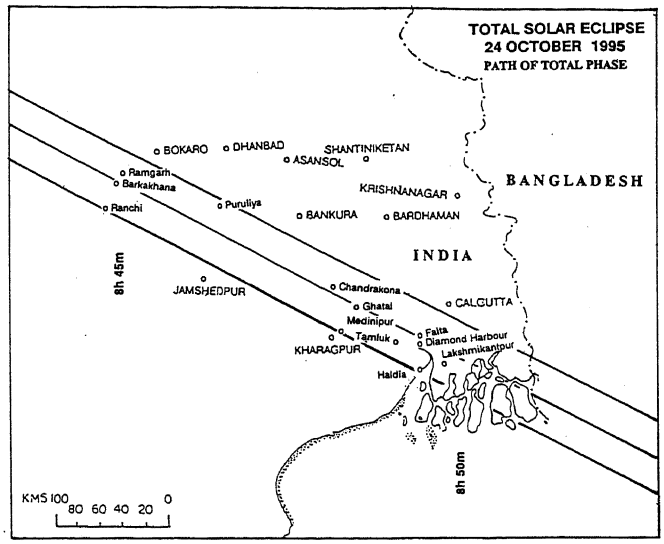
For this year's eclipse, detailed maps showing the eclipse path in India have been given by Singh[3]. I am giving only some important points about the event from the viewpoint of observers from India. The first glimpse of totality will be from a village Anandgarh on the Indo-Pak border at 0830 IST, with the sun 20° above the horizon; its duration will be about 45 s. Two fairly big towns in Rajasthan, Alwar and Bharatpur will fall in the eclipse path though not on the centre, as a result of which the duration of totality for viewers located at these places will be reduced. The cities of Agra and Etawah will fall just outside the path, while Allahabad and Mirzapur will be just inside; but to get a good view of the eclipse the viewers will have to move some 10–15 km to reach the centre of the eclipse path. In Bihar, both Ranchi and Hazaribag will fall a few km outside, so also the metropolis of Calcutta, whose southern suburban area will be just inside the eclipse path. Duration here on the centre of the track will be about 80 s, with the sun's altitude about 40° at mid totality.

After this year's eclipse, one more eclipse track will cross over India in the late afternoon of 11 August 1999. The course of the shadow will cross Gujarat, Madhya Pradesh and Andhra Pradesh and will end at sunset over the Bay of Bengal. But considering the seasonal chances of monsoon clouding over the eclipse path and the lower altitude of the sun in that event, this year's eclipse appears to be the last good chance for Indian scientists in this century for arranging experiments which will hopefully produce some vital clues to the mystery of the sun.

The scientific explanations for the occurrences of total eclipses and various observable events are available in the literature; scientific experiments undertaken during such events are also extensively covered in some review articles[4]. We may skip those points to avoid unnecessary repetition and concentrate on various features observable by amateur astronomers and general viewers. We may carefully note one warning: looking at the bright disc of the sun without protection is hazardous. During an eclipse event, most of the time the uneclipsed part will be visible, it may only be viewed through a heavily darkened filter. The type of goggles which the welders use are quite suitable; they cut down the brightness to about 10^{-4}, which is essential. An alternative can be prepared by exposing black and white photographic films to daylight (not direct sun) and developing them fully in the dark room. The films can then be fixed in hypo, washed and dried in air. One layer may not provide the extinction needed; two or three layers mounted together may be necessary. The sun through this composite may look brownish and can be viewed without any discomfort.

Once armed with this device, we can venture to view an eclipse event. Let us imagine that we are located at the centre of the path of totality. The invisible moon will be approaching the sun from the west; the first contact will appear as





an almost unnoticeable dent in the solar disc. Minute by minute the eclipse will grow; slowly we will be able to recognize the dark shape of the moon. The moon appears here as pitch dark, even though considerable light reflected from the earth illuminates its face. This appearance is purely due to contrast to the uneclipsed part of the bright photosphere.

Not very spectacular things seem to have happened as yet: there is a gradual decrease in general illumination and the air temperature starts falling, and this can be felt. I have been told that confusion can be noticed among flocks of birds and herds of cattle at the untimely appearance of the dusk, but am unable to describe in proper sequence from my own experience, because whenever I faced such a situation in the past, I was much too occupied and worried about my experiments to pay attention to these details. But one does not miss some amount of eerieness in the atmosphere at that time.

Something noticeable happens in the shadows cast by obstacles. If we have a tree nearby then through some gaps in the foliage tiny bits of sunlight fall on the ground. In normal times these take the form of diffuse circles of light but now one may clearly notice sickle-shaped light spots instead. In fact, these are pinhole images of the partially eclipsed sun through the narrow openings in foliage. It is as expected from the laws of physics, but nevertheless quite striking in appearance.

Now let us turn our attention to the sun; the sun is already in a crescent form, something similar to the crescent moon, only the edges are much sharper. They get thinner and thinner and ultimately break up into a string of bright beads; these are called Baily's beads after the scientist who first described the phenomenon in 1836. The appearance is due to unevenness of the lunar limb; the photosphere may shine through small valleys for a few seconds more after almost the entire disc has been covered by the moon. Sometimes the presence of a deep valley may produce a 'diamond ring' event, when just one very bright point as if mounted on a ring of pink chromosphere may be seen. Any idea of the beauty of this diamond ring is difficult to convey to a person who has not seen it.

One more interesting piece of observation may be made at this juncture; bands of light and shadow may be seen on the ground, rippling all over. There are different appearances of these shadow bands, and in that unreal atmosphere the observers imagine different things. Some see them as if standing in a snake-pit with vipers crawling all around, and some see the event as faint shadow ripples passing under their feet. The phenomenon is similar to the twinkling of stars and owes its origin to the irregular refraction by atmospheric eddies. The only difference is that shadow bands cast in the light of bright stars are not detectable in reflected light, but the tiny Baily's beads are much brighter and the bands created in their light are easily detected by unaided eye.

The pink chromosphere flashes out as soon as the photosphere is covered; if the sun is active, a few prominences, plumes of glowing solar gas stretching out of the limb may be seen. The view lasts for one or two seconds only, but this short stretch of time is most eagerly awaited by solar physicists, ever since Janssen and Tennant did the pioneering experiments on the tobacco fields of Guntur on the morning of 18 August 1868, when a new element helium was discovered, 27 years before it was detected in a British laboratory by William Ramsey.

For amateurs and general eclipse viewers I may safely recommend an interesting observation. The only instrument one will need is a pocket-sized direct vision spectroscope. One should keep looking through this instrument at the thinning crescent sun; the protective dark filter should be removed; the spectroscope will disperse sunlight automatically bringing down its intensity to a safe level. One will see dark Fraunhofer lines clearly; the moment the photosphere is covered by the moon, all dark lines will flash into an array of vividly coloured emission lines. This experiment was first performed by C. A. Young of Princeton University during the eclipse of 22 December 1870, and ever since it has been repeated almost in all eclipses, not only for the valuable information about the detailed structure

of the solar chromosphere but also for the sheer magical beauty of the event.

As the moon will cover up the thin chromosphere, the pink light will fade and be replaced by the pearly white glow of the solar corona. I must repeat that the protective dark filter will have to be removed during totality, otherwise all of these interesting views will be missed. The size and shape of the corona varies from eclipse to eclipse, depending upon the phase of the solar activity cycle. At sunspot maximum the corona will be extensive and covers all sides, while at minimum it generally avoids polar regions and extends mainly over the solar equator. This year's eclipse will be near a sunspot minimum and the latter shape is expected. But even then, the glow of the corona will prevent the sky from becoming totally dark; there will be almost as much light as that from a full moon, but it will be more mysterious, more ethereal, spinning a web of fantasy over the whole atmosphere.

The pictures of the solar corona we see on the pages of astronomy textbooks are far from inspiring, most of the intricate beauty being lost in the recording process. But the eye can perceive many details of streamers and rays, and catch the subtle display of light and its variations over the precious seconds of totality. Of course, the new recently developed techniques of scientific processing of the coronal photographs in the laboratary can bring out part of the lost information, though not the beauty of the grandest spectacle of nature.

The totality ends rather suddenly, with a brief appearance of chromospheric glow, Baily's beads and, if lucky, with a diamond ring event; then one can see even the moon's shadow racing away, and daylight reappearing like a flood. The eclipse is over, although the crescent sun will take almost an hour more before assuming the normal full disc form.

It is indeed a pity that many persons refuse to see the total eclipse because of superstitious ideas. Some think that harmful radiations emanate from the sun during eclipses; some even go to the extreme of imagining that poisonous gases pervade the earth's atmosphere. The unusual happenings in the sky used to strike terror in the minds of primitive men, who imagined demons devouring the sun and eventually destroying everything; the customary penances and ritual baths after the event are remnants of the original feeling of terror. The ideas of a demon devouring the sun have largely disappeared, but the fear of the unknown persists. I know persons who not only refused to observe the eclipse, but kept their entire families locked indoors to prevent them coming under the influence of the eclipse; even tiny holes in their windows were plugged to stop the light from the eclipsed sun to enter their homes. It is difficult to explain to them that nothing happens to the sun during eclipses; the shadow cone cast by the moon is always floating in space, and this just goes over the earth's surface during eclipses. Even the idea of generation of poisonous gases or bacteria in the shadow cone is unscientific; the earth also carries its own shadow cone on the night side, everyday we are living in it for twelve hours. It is futile to point out that these ideas are totally based on ignorance and have no trace of scientific logic. The persistence of the feeling is aided by the extreme rarity of the phenomenon; for people rooted to their homestead, the chance of viewing a total eclipse comes once in fourteen generations.

Many point out the occurrence of eclipse blindness which had totally spoilt the eyesight of many persons in the past. This fear is genuine but need not be attributed to any strange unnatural phenomena during eclipses. The real reason lies in the extreme brightness of the sun; if any person keeps on looking at the sun, his retina will also get burnt out, causing blindness. In normal times physiological reflexes prevent such mishaps but during the penultimate minutes just prior to totality, general illumination falls and human iris opens up to collect as much light as possible in the rapidly darkening surroundings. But the crescent sun retains the same intrinsic brightness per unit area, and its image focused on the retina with the fully opened eye-lens burns out a sickle-shaped area of the sensitive tissue. The dangerous aspect of this accident is that the observer does not feel any pain, and the damage is done without the viewer's knowledge. Prevention of such happenings is simple and has been prescribed earlier in this article. Nothing happens to the skin or any other part of the body by exposure to the eclipsed sun.

To conclude, I may say that this is a rare opportunity for many to witness an event which I may call as Nature's grandest spectacle. I am still to meet any person who has gone and witnessed a total solar eclipse, and repented the trouble and extra effort needed for his adventure.

On the eve of the total solar eclipse of 16 February 1980, Dr M. K. V. Bappu had addressed the Osmania University faculty and the students; I feel his concluding advice may be retold now. An excerpt below gives beautifully what I wish to convey to all enthusiastic viewers:

'I now think of the man who enjoys some of the nice sights of Nature, one who has got an aesthetic sense of feeling of appreciation of beauty, who enjoys a good sunset, who finds a feeling of delight at hearing the roaring waters of a waterfall or of a gurgling stream as it passes by, who delights in a good rainbow or takes a fancy for seeing the well-spun cobweb of a spider on a dewy morning; in short, an individual who is alert to his environment and who enjoys different aspects of it. I think there is much for him in witnessing one of these rare sights of the total eclipse of the Sun. I hope that he would just travel down not very far, just even twenty miles south of Hyderabad and take with him, probably a minimum of equipment, just his own eyes, and go through this delectable feast that is really only a feast of the Gods.'[5]

I have only to substitute the city of Hyderabad by a chain of cities in North India for this year's eclipse; otherwise, the advice remains the same for millions who do not wish to miss an opportunity of a lifetime.

1. Espenak, F. and Anderson, J., NASA Reference Publication No.1344, 1993.
2. Singh, J., *Bull. Astron. Soc. India*, 1994, **22**, 339–354.
3. Singh, J., *Bull. Astron. Soc. India*, 1995, **23**, 3–12.
4. Bhattacharyya, J. C., *Indian J. Radio Space Phys.*, 1990, **19**, 525–533.
5. Bappu, M. K. V., Contr. Nizamia & Japal Rangapur Observatories Contribution, No.10, 1979.

J. C. Bhattacharyya, Indian Institute of Astrophysics, Bangalore, India.

10th Himalaya–Karakoram–Tibet Workshop, Switzerland, 1995 – Conference report

Since 1985, a series of international annual meetings on the geology of the Himalaya–Karakoram–Tibet region have been held: Leicester, UK (1985); Nancy, France (1986); London, UK (1987); Lausanne, Switzerland (1988); Milan, Italy (1990); Grenoble, France (1991); Oxford, UK (1992); Vienna, Austria (1993); and Kathmandu, Nepal (1994). The 10th Himalaya–Karakoram–Tibet Workshop was held at Monte Verità, Ascona city, in Switzerland on April 3–8 1995. This was the second time that the Swiss held the Workshop, and it was organized by David Spencer, Jean-Pierre Burg and Cinzia Cervato-Spencer of the Institute of Geology, Swiss Federal Institute of Technology (ETH), Zürich. A combination of several factors made this meeting very successful. The ETH geologists have a long record of research in the Himalaya. Augusto Gansser (now 85) is a pioneer Himalayan geologist. His book *Geology of the Himalayas*[1] is a cornerstone of our geologic knowledge of the Himalaya. Gansser, former Chairman of the ETH Institute of Geology and now a retired Emeritus Professor, opened the Workshop with a fascinating talk on his 60 years of geologic work in various parts of the Himalaya, including his visit to the holy mountain of Kailash in 1936. Gansser and Arnold Heim (the son of the famous alpine geologist Albert Heim) published the results of their field work in Darjeeling, Kumaun and south Tibet in a treatise, *Central Himalaya*[2], an original and pioneering piece of work.

The Workshop was well organized. It was held in a charming spot in Switzerland: the ETH Conference Centre at the Alpine hilly station of Monte Verità facing the blue lake of Maggiore. Fees for accommodation and food were subsidized. And more important, 25 participants came from the Himalayan countries – thanks to the fund-raising efforts of the conveners. Some 144 participants from 17 different countries in Europe, North America and Asia attended the Workshop.

The Himalayan Workshops are a timely response to increasing interest among geologists in the Himalaya and Tibet. With an area of some 2.5 million km^2, a mean elevation of 5000 m above sea level and crustal thickness of about 70 km, the Tibetan Plateau is the world's largest highland. Its geological genesis is still an enigma. The Himalaya bordering the southern margin of Tibet is the loftiest and youngest mountain chain on earth. These spectacular mountains hold crucial keys to understanding how mountains resulting from the collison of continental plates form. They are also linked in many ways to the ecology and life in south Asia. Rivers, rain-bearing monsoon clouds, and sediments coming from the Himalaya support agricultural communities, accounting for more than 10% of the world's population; they are also responsible for the formation of the world's largest delta (the Ganges Delta) and the largest marine fan (the Bengal Fan). In recent years, it has also been suggested by Raymo *et al.*[3] that the uplift and erosion of high mountains, especially the Himalaya–Tibet region, initiated the ice ages in the northern hemisphere during the past 3 million years. This is possible because high rates of chemical weathering and denudation absorb atmospheric carbon dioxide. This hypothesis was recently featured in a BBC documentary entitled 'Tibet: The Ice Mother', which was also shown at the Workshop.

It was, therefore, appropriate that a special session of the Workshop (with 10 oral presentations) be devoted to 'Geological processes related to uplift, exhumation and elevation of the Himalaya, Karakoram and Tibet'. This author began the session by presenting a sequential history of the uplift and denudation of the Himalaya throughout the Cenozoic, and emphasizing that the rise of the Himalaya seems to have occurred episodically in several distinct phases rather than a steady-rate slow history, a single big boom or merely chaotic events here and there. A one-hour Discussion Meeting had also been scheduled to discuss informally 'The mechanisms and consequences of Tibetan plateau uplift', led by Peter Molnar, a geophysicist who has been pondering on the geodynamic evolution of the Himalaya and Tibet since the early 1970s (Molnar and Tapponnier's influential paper[4] in *Science* is a must to read). During discussions on the uplift history, it seemed that two major ideological camps among Himalayan researchers exist: those who confine the uplift of the Himalaya and Tibet to the Miocene epoch (23–5 Ma) (e.g. Molnar), and those who also add a major Quaternary (the past 1.6 million years) phase of uplift to this history (e.g. Gansser).

We know that the Himalaya grew out of the closure of the Tethyan Ocean, which once lay between India and Asia. We also know that the initial boundary between the Indian and Asian tectonic plates, along which Tethys was consumed and subsequent India–Asia collison took place, is represented by the Indus–Tsangpo Suture Zone, first suggested by Gansser in 1964. But the timing of this collision is controversial. Richard Beck of the University of South California presented a summary of his research group's recent paper[5] in *Nature* presenting stratigraphic evidence from the Indus Suture Zone in Pakistan that the collision took place between 65 and 57 million years ago (Ma) at least in northwestern Himalaya. Previously, Searle *et al.*[6] had argued that the collison was at 50 Ma. Another one-hour discussion was devoted to 'The timing of the India–Eurasia collision', headed aptly by Searle and Beck. Much stimulating discussion followed, and with a mood of 'back to the basics', the very definition of 'continental collision' was questioned. 'Continental collision' goes through an early stage of 'ocean floor termination', i.e. when the basaltic ocean floor is completely consumed by the process of subduction, and a later 'continental suturing', i.e. when the shallow marine facies changes into purely continental sedimentation. Beck defined the 'collision' as the former, and Searle as the latter. But why not consider both of them as initial and end

members of 'continental collision'? – ocean floor termination initiating the deformation of the colliding continents, prominently involving the accretionary prism, although shallow marine sedimentation is still going on, and continental suturing giving rise to land conditions at the site of the collision. There may be a time span of several million years between these two events. In other words, drifting and converging continental plates 'hit' each other before they actually 'meet' each other.

The Himalaya as we know consists of a series of fold-and-thrust belts. The backbone of these mountains is the Central Crystalline Zone of the Higher Himalaya, which has thrust over the Lesser Himalayan sediments and low-grade metamorphosed sediments along an orogen-scale rupture known as the Main Central Thrust (MCT). Some 20 years ago, Le Fort[7] proposed a model according to which simultaneous metamorphism and thrusting along the MCT brought 'hot' rocks of Higher Himalaya over the 'cold' Lesser Himalaya, causing an 'inverted metamorphism', i.e. metamorphism temperature–pressure conditions increasing upwards the mountains. This model was put forward to answer the enigma of inverted metamorphism in the Himalaya, first observed by several geologists in the late 19th century. There are, however, several other models to explain inverted metamorphism, including granitic intrusions in the Higher Himalaya, differential thrusting of each metamorphic zone, large-scale folding of metamorphic isograds, and ductile shearing distributed across the Central Crystalline rocks. Several talks addressed these issues, and a one-hour Discussion Meeting was devoted to '20 years of MCT models – Which ones work?'. It seems that to relate the Himalayan inverted metamorphism to the MCT a clear definition of the MCT along the strike of the Himalaya is required (Mr. X's MCT in Garhwal may be entirely different from Mr. Y's MCT in Nepal and hence confusing debates) and then it should be considered whether the MCT is a unique single structure in terms of Himalayan metamorphic history or rather a wide zone of multiple thrusts and shearing. Systematic thermochronological and thermobarometric data across the Higher Himalaya are also required to address the 'puzzle' of inverted metamorphism. In the end, we may find that each model, like the four blind men in the famous Indian parable, touches a certain aspect of 'the elephant' and conveys some but not all of the truth, of course with all the likeliness that some of the models are more truthful than others.

John Ramsay, a well-known structural geologist who succeeded Gansser at the ETH Institute of Geology and is now retired (himself succeeded by Jean-Pierre Burg, a convener of the Workshop) gave an overview of the discussions during the four-day Workshop. Ramsay remarked that as Emile Argand's 1916 map of the western Alps has changed little after eight decades, Gansser's 1964 geological map of the Himalaya is essentially valid to this day. What aspects of Himalayan geology may be expected to yield novel discoveries in the future? From the standpoint of structural geology, Ramsay emphasized the need to study the deformation history within the geological divisions of the Himalaya rather than merely viewing each orogen-scale division as an intact unit.

The Abstract Volume was published as No. 298 issue of *Mitteilungen aus dem Geologischen Institute*, ETH (349 pp). A proceedings volume of the Workshop containing selected, referred papers is scheduled to be published in *Tectonophysics* in 1996. A total of 48 oral presentations were made in 10 sessions: Tectonics and Palaeoclimate (2 talks), Southern Tibet (4), Tibetan Plateau (3), Himalayas of Pakistan and India (14), the Main Central Thrust (4), Central and Eastern Himalaya (14), Asian tectonics (4), and Applications of geochronology in the Himalaya (3). Some 52 posters were displayed and each poster was given three minutes to convey its main message, which was a good strategy to handle poster presentations. Overall, the coverage of topics was well balanced, although at the end of the Workshop there was criticism that usually in the Himalayan Workshops much attention is paid to 'hard rock' geology, while many important clues to understand the evolution of the Himalaya are hidden in its sedimentary record ('soft rock' geology). Interestingly, this criticism was made by two stratigraphers, Aymond Baud (Switzerland) and aurizio Gaetani (Italy) as well as by a tectonist, Vikram Thakur (India).

Most of the Indian participants at the Workshop came from the Wadia Institute of Himalayan Geology (Dehra Dun) and the Birbal Sahni Institute of Palaeobotany (Lucknow), which are, in my opinion, two of the best geoscientific research institutions in India, aptly named after the great Indian scientists: D. N. Wadia (1883–1969) and Birbal Sahni (1891–1949). India with its vast resources in Himalayan geology should try to hold one of the Himalayan Workshops in the future – before the new millennium begins!

1. Gansser, A., *Geology of the Himalayas*, Interscience, London, 1964, pp. 289.
2. Heim, A. and Gansser, A., *Mem. Soc. Helv. Sci. Nat.*, 1939, **73**(1), 1–245. (Reprinted by Hindustan Publ. Corp., Delhi, 1975).
3. Raymo, M., Ruddiman, W. F. and Froelich, N. J., *Geology*, 1989, **16**, 649–653.
4. Molnar, P. and Tapponnier, P., *Science*, 1975, **189**, 419–426.
5. Beck, R. A., *et al.*, *Nature*, 1995, **373**, 55–58.
6. Searle, M. P., *et al.*, *Geol. Soc. Am. Bull.*, 1987, **98**, 678–701.
7. Le Fort, P., *Am. J. Sci.*, 1975, **A275**, 1–44.

Rasoul B. Sorkhabi, Department of Geology, Arizona State University, USA.

A good beginning

Current Science deserves appreciation for the courage and the cudgels it has taken up to improve the standards of teaching science and scientific research in our universities. There is no gainsaying the fact that owing to several articles published in *Current Science* in the last two years or so the subject of deteriorating quality of science, both teaching and research, in our universities is no more a sacred cow. Several conscientious academicians have voiced their opinion that the onus for this unsatisfactory state of affairs, several things apart, lies to a fair degree of measure on the scientific and teaching community also[1,2]. It is heartening to see that in terms of objectivity and critical appraisal of scientific research being conducted in our universities and research institutions, the articles published in *Current Science* score over those which appeared sometime back in *Nature*[3], in which every thing about scientific research in India was goody goody and the other side was totally overlooked. Lack of critical self-appraisal, which, of late, has come to pervade our scientific and educational institutions of higher learning is one of the important causes of our mediocrity in these areas. It is high time that we get ourselves rid of this malady lest it is passed on to our future generation of teachers and scientists. In this regard we need to take a lesson from S. Chandrasekhar, whose recounting of the story of Daedalus and Icarus elegantly brings out how critical self-appraisal is fundamental to all good research and advancement of science[4].

The initiative of the Council of the Indian Academy of Sciences to improve the standards of science education in the country is timely and the agenda enunciated in the 'Academy Paper'[5] is both laudable and challenging. If the articles published in *Current Science* have struck a chord in the hearts of our young researchers and teachers then, probably, a good beginning has already been made.

1. Balaram, P., *Curr. Sci.*, 1994, **67**, 502–503.
2. Srivastava, P. N., *Curr. Sci.*, 1994, **67**, 508–512.
3. Maddox, J., *Nature*, 1993, **366**, 611–626.
4. Chandrasekhar, S., *Truth and Beauty. Aesthetics and Motivation in Science*, Penguin, New Delhi, 1991.
5. University Education in Science, *Curr. Sci.*, 1995, **68**, 255–267.

J. S. VIRDI

Department of Microbiology,
University of Delhi South Campus,
Benito Juarez Road,
New Delhi 110 021, India.

Does low serum cholesterol increase suicide and aggressive behaviour?

Amar Bavle

It is a well-established medical fact that high serum cholesterol is a risk factor for developing ischaemic heart disease and that lowering the high level of serum cholesterol in these patients reduces mortality. Interestingly, follow-up studies showed no overall reduction in the mortality rates. A meta analysis of six primary prevention trials of cholesterol-lowering agents in men (pooled $N = 24,847$) by Muldoon *et al.*[1] revealed that the intervention group experienced reduction of deaths from heart disease compared to controls but the total mortality remained the same. This means that with reduction of serum cholesterol, mortality from other causes increases.

This prompted a detailed examination of the causes of noncardiac mortality. Epidemiological studies have examined the relationship between nonmedical mortality and baseline low serum cholesterol in the general population. The Varmland study in Sweden by Lindberg *et al.*[2] and the study of Neaton *et al.*[3], both in 1992, screened more than 350,000 men and showed a higher mortality in men with low serum cholesterol. Most unexpectedly, the higher mortality from noncardiac causes was found to be mainly due to suicides and traffic accidents and aggression as well. A higher risk of suicide and deaths from accidents were reported from a meta analysis that pooled the results of 18 prospective epidemiological studies[4].

More recent findings[5] suggest that men with serum cholesterol levels at or below the 25th percentile were twice as likely to have ever made a serious suicidal attempt compared to men whose serum cholesterol was above the 25th percentile. They have proposed that low serum cholesterol is a potential biological marker of suicidal risk.

Is it possible to explain the increased suicide rates in people with low serum cholesterol? There is an interesting hypothesis to explain this. The explanation invokes the serotonin hypothesis of depression, suicide and impulse control disorders (which includes aggressive behaviour). All these conditions are associated with low serotonin levels in the brain and low serotonin levels in the human brain are associated with poorer suppression of behavioural impulses

like suicide and aggressive behaviour[6].

There is evidence that low serum cholesterol is associated with a number of aggressive behaviours. For example, varied psychiatric disorders where aggression is an important symptom, like in male homicidal offenders who are habitually violent under the influence of alcohol[7], and male criminals with antisocial personality disorder who are aggressive[8]. Even from animal experiments there is some evidence that a cholesterol-reducing diet in monkeys makes them more aggressive[9]. A low cholesterol diet in monkeys is also associated with low brain serotonin levels[10]. Both suicide and aggressive behaviour, therefore, are associated with lower serum cholesterol possibly reducing brain serotonin levels[6].

In summary, there is evidence that noncardiac deaths, particularly suicides, occur not only in patients receiving cholesterol-lowering drugs but also in people whose serum cholesterol is naturally low. Interestingly, for some unknown reason, this effect is much more pronounced in men than in women. The question that is yet to be answered is: Is it correct advice to tell people to keep their cholesterol levels low? Cardiologists will invariably say yes, because it reduces mortality from cardiac conditions. It would be wiser to review the answer in the light of recent findings and look at it from all angles rather than just the narrow cardiac viewpoint.

It would be interesting to examine to what extent these patterns are seen in the Indian population. These investigations have recently been undertaken.

1. Muldoon, M. F., *et al.*, *Br. Med. J.*, 1990, **301**, 309–314.
2. Lindberg, G. *et al.*, *Br. Med. J.*, 1992, **305**, 277–279.
3. Neaton, D. *et al.*, *Arch. Int. Med.*, 1992, **152**, 1490–1500.
4. Jacobs, D. *et al.*, *Circulation*, 1992, **86**, 1046–1060.
5. Golier, J. A. *et al.*, *Am. J. Psychiatr.*, 1995, **152**, 419–423.
6. Engelberg, H., *Lancet*, 1992, **339**, 727–729.
7. Virkhunen, M., *Neuropsychobiology*, 1983, **10**, 65–69.
8. Virkhunen, M., *Neuropsychobiology*, 1979, **5**, 27–30.
9. Kaplan, J. R., *et al.*, *Psychosom. Med.*, 1991, **53**, 634–642.
10. Muldoon, M. F. *et al.*, *Biol. Psychiatr.*, 1992, **31**, 739–742.

ACKNOWLEDGEMENT. I acknowledge the critical comments, suggestions and encouragement offered by Dr N. V. Joshi in preparing this paper.

Amar Bavle is in the Department of Psychiatry, M. S. Ramaiah Medical College, Bangalore 560 054, India.

Bose–Einstein condensation in a dilute atomic vapour

N. Kumar

A team of researchers working at the Joint Institute for Laboratory Astrophysics (JILA) in Boulder, Colorado, has just announced[1] having observed three distinct signatures for the occurrence of Bose–Einstein condensation (BEC) in a dilute ultracold vapour of atomic rubidium-87 (^{87}Rb). The most distinctive signal claimed for the onset of the BEC was an abrupt appearance (and gradual growth) at (and below) a threshold temperature of an anisotropic momentum distribution sharply peaked about zero against a diffuse, isotropic thermal background that reflected the lowest-energy *single*-particle state, now being occupied *macroscopically* (i.e. extensively) by the bosonic ^{87}Rb atoms. And that is precisely what the Bose–Einstein condensation is all about – a *macroscopic* (*finite* fraction of the total) number of particles occupying the lowest-energy *single*-particle state, leaving the higher states sparsely populated.

The experiment involved (i) confining a sample of about 2000 ^{87}Rb atoms to a volume about 10 μm across, thus compressing the vapour to a number density of about 2.6×10^{12} cm^{-3}, and (ii) cooling it to an abysmally low temperature of about 170 nk and maintaining the BEC so obtained for a reasonable length of time of about 15 s for diagnostic studies. All this necessitated a synergetic combination of novel experimental techniques for trapping and cooling of atoms. Thus, the magnetic trap used was essentially a quadrupolar magnetic field that dotted with the antiparallel-aligned magnetic moments of ^{87}Rb atoms ($F = 2$, $m_F = 2$) so as to give a confining potential well. This potential well, however, had at its central minimum the magnetic field equal to zero, allowing the atoms to flip their moments freely through unavoidable perturbations and escape. This *leak* was *plugged* in an ingenious manner by superimposing a rotating magnetic field, slow enough for the moments to follow adiabatically but fast enough to yield a time-averaged orbiting potential (TOP) which was parabolic, and with a non-zero averaged magnetic field at the minimum that prevented the spin flip. The TOP was a uniaxial 3D harmonic potential giving an oblate ground state wavefunction. The latter was reflected in the anisotropic momentum (velocity) distribution of the BEC recorded by shadow imaging. As for the cooling, the atomic gas was precooled by the laser Doppler technique, in which the atoms are retarded by a set of counterpropagating laser beams with their frequency tuned slightly below that for an atomic resonant absorption. The Doppler shift then ensures that the moving atom has the sisyphean task of always having to climb up a potential hill and hence get retarded to a near-zero speed. This is then followed by a *forced evaporative cooling* in which the more-than-average energetic atoms are allowed to escape the magnetic trap at the edges, leaving the tardier atoms to thermalize to a lower temperature. The evaporation is ingeniously *forced* by an rf magnetic field that flips the spins of the

energetic atoms near the edge. This instantly inverts the potential well, thus facilitating escape. These techniques, developed and perfected over the years building on the work of several research groups, enabled the researchers to create finally the extreme condition of phase space density $n\lambda_T^3 = 2.621$ required for BEC. Here n is the number density and $\lambda_T = h/(2\pi m k_B T)^{1/2}$ is the thermal de Broglie wavelength of the condensing atoms.

Now, the idea of Bose–Einstein condensation is by itself not new, having been predicted by Einstein in 1925 for ^{4}He, basing on the then new quantum Bose statistics for particles with integral spin, such as photons, proposed by Satyendra Nath Bose[2]. Unlike their half-integral 'exclusive' cousins, the fermions (such as electrons or ^{3}He), bosons are gregarious, and arbitrarily large number of them are allowed (in fact, are *encouraged* to) occupy a given single-particle state. At a sufficiently low (high) temperature (density) this creates an overpopulation crisis and a macroscopic (finite fraction of the total) number condenses into the lowest single-particle state. This, of course, happens only for permanent bosons with a fixed given number (e.g. ^{4}He) and not for photons, say, where the number keeps diminishing as the temperature is lowered and hence no BEC. At this phase-space density, the de Broglie wavelength is comparable to the mean interparticle spacing (overlapping wavefunctions), and the quantum-mechanical indistinguishability of identical particles asserts itself by making the condensate behave as a single coherent entity. Stable BEC is known to exist in liquid ^{4}He below 2.2 K, where it is implicated in superfluidity. It has been achieved transiently in a gas of excitons, electron-holes ($e^- - e^+$) bound pairs created optically in a semiconductor like cuprous oxide[3]. Spin-polarized hydrogen (H↓) has been for the last 15 years a strong candidate: it remains gaseous down to absolute zero of temperature[4]. Then why so much excitement about BEC in ^{87}Rb, one may ask. Of course, the threshold phase-space density for BEC is easier to obtain for the lighter atoms. But, the purely quantum-statistical effects are corrupted by the strongly interacting nature of these dense systems. By comparison, for the dilute ^{87}Rb gas in question, the mean interparticle spacing ($\sim 10^{-4}$ cm) >> the scattering length ($\sim 10^{-6}$ cm) and hence the near-ideal Bose gas condition prevails. (The magnetically trapped gas is metastable against crystallization at nanokelvin temperatures). Also, ^{87}Rb with conveniently accessible electronic transitions is manipulable, even its interactions are optically tunable. One can probe the entire quantum phase diagram – dilute gaseous BEC, dense liquid BEC, superfluidity, quantum solid and possibly superfluid solid! Also, the kinetics of the Bose–Einstein condensation itself[1]. Besides, compared to the aristocratic quantum fluids, e.g. ^{4}He, the rubidium-87 (an alkali) is really ordinary!

In what sense is BEC *coherent matter*? This is best illustrated by its possible interaction with light, e.g. photoassociation[5]. For a condensate of N atoms, the optical transition-dipole will be $\sqrt{N}$ times that for a single atom even if the condensate size is less than the wavelength of light. Coherent matter analogue of optical laser is also being speculated upon!

One could envisage several other basic studies – Anderson localization (and confinement) of BEC by random potential created by an optical speckle pattern simulating quenched disorder. One could also examine the fundamental question of the effect of frictional (γ) decoherence on BEC – one may not have any BEC at all for $\gamma > (\hbar/m)\, n^{2/3}$ even at the absolute zero of temperature. The friction γ can be tuned optically, or by dilution with noncondensing (fermionic) atoms. One expects a BEC to have a negative temperature coefficient of expansion. And so on.

One may conclude by saying that if phase and coherence are to dominate our thinking in the years to come, then the coherent matter (BEC) may well be a laboratory of choice.

1. Anderson, M. H., Ensher, J. R., Matthews, M. R., Wieman, C. E. and Cornell, E. A, *Science*, 1995, **269**, 198.
2. Venkataraman, G., *Bose and His Statistics*, University Press, Hyderabad, 1992.
3. Lin, J. L. and Wolfe, J. P., *Phys. Rev. Lett.*, 1993, **71**, 1222.
4. Greytak, T. J., in *Bose–Einstein Condensation* (eds Griffin, A., Snoke, D. W. and Stringari, S.), Cambridge University Press, Cambridge, 1995.
5. Kleppner, D., *Phys. Today*, July 1995, p.11.

N. Kumar is in the Raman Research Institute, C. V. Raman Avenue, Bangalore 560 080, India.

SCIENTIFIC CORRESPONDENCE

'Swiss' strain of rats?

The strain of rats used in the study (A. K. Mitra *et al.*, *Curr. Sci.*, 1995, **68**, 1050–1053) is given as 'Swiss strain'. As one representing an organization concerned with laboratory animals maintenance and supply, I can authentically say that 'Swiss' strain of rats does not exist. Most of the laboratory rats currently available for biomedical research have originated from the brown Norway rat – *Rattus norvegicus* and some common strains originating from this are Wistar, Holtzman, CFY and Sprague–Dawley to mention a few. But most of the mice strains have a 'Swiss' origin and many random-bred albino mice available in various laboratories are loosely termed as 'Swiss' strain. Obviously there is a mistake somewhere and it is better that the authors clarify this from their 'aunthentic breeders'.

N. GIRIDHARAN

Laboratory Animals Information Centre, National Institute of Nutrition, Jamia-Osmania PO, Hyderabad 500 007, India.

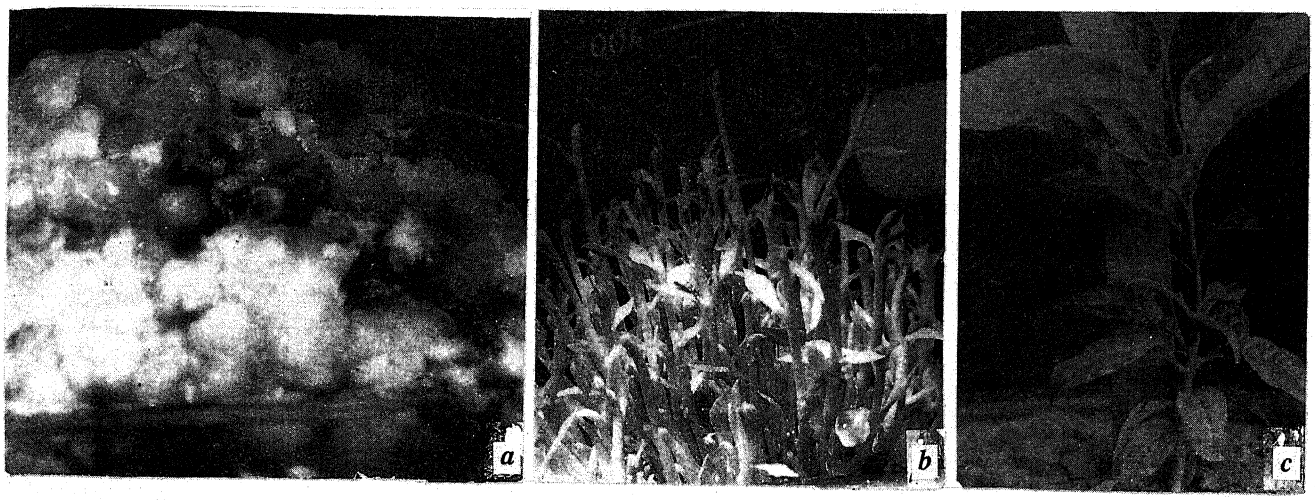
Regeneration of *Aegle marmelos* (L.) Corr. plants *in vitro* from callus cultures of embryonic tissues

Aegle marmelos is a medium, spreading deciduous spiny tree belonging to the family Rutaceae. It is supposed to be originated in the Indo-Burma subcontinent[1]. Various parts, including fruits, possess medicinal properties and have been extensively used in ayurvedic and folk medicine[2,3]. Since *A. marmelos* is an outbreeding crop and cross-pollination leads to highly heterozygous population[4], propagation through seeds is not encouraged. Singh[5] was first to culture *A. marmelos* but regeneration of plants from nucellar tissue was not possible. Arya *et al.*[4] described callus formation and some organogenesis from cotyledon and hypocotyl cultures. Recently, regeneration of multiple shoots through organogenesis was achieved from seedling leaves[6,7], nucellus[8] and cotyledons[9]. This communication presents high-frequency plant regeneration system using callus cultures from zygotic embryos of *A. marmelos*.

Mature seeds from ripe fruits and immature seeds from unripe green fruits were collected and surface-sterilized with 0.1% $HgCl_2$. Seeds were decoated and embryo axes excluding cotyledons were excised and cultured. All cultures were maintained in a growth room under 16 h photoperiod at 26°C. All media were solidified with 7 g/l agar (BDH, high gel strength) and media pH values were adjusted to 5.7 ± 0.1 prior to autoclaving. The basic culture medium was Murashige and Skoog (MS) salts[10] variously supplemented with naphthalene acetic acid (NAA), 2,4-D, IAA, IBA, BA and kinetin. Different concentrations of auxins in the range of 2.5–10 µM were used for callus induction. Plant regeneration was achieved by transferring pieces of callus tissue (fresh weight 100–150 mg) on to the MS medium with different combinations of auxins and cytokinins. Rooting was induced in IBA supplemented (5 µM) MS medium. After rooting, the leafy shoots were washed and transferred to pots containing soil mixed with compost (2:1).

Different concentrations 2.5–10 µM of IAA, NAA and 2,4-D were tried to induce callus and it was observed that 2,4-D in general was more efficient for the induction of callus and induced as much as 80% callus from both mature and immature embryo axes at a concentration of 10 µM. However, addition of a lower concentration (0.5 µM) of BA together with 10 µM 2,4-D induced 100% callus formation in all explants obtained from mature and immature seeds. The 6-week-old hard, green and compact irregularly shaped calli (Figure 1 *a*) induced on 10 µM 2,4-D and 0.5 µM BA were used to study morphogenesis. Such calli obtained from each explant were cut into small pieces (100–150 mg fresh weight) and were transferred to fresh medium. Both BA and kinetin at various concentrations were capable of inducing differentiation from undifferentiated calli (Table 1). Of the two cytokinins, kinetin was better than BA and of the auxins IAA were more efficient than NAA. Thus, the best results for shoot formation were obtained using kinetin at 5–10 µM and IAA at 0.5 µM. In all cases both mature- and immature-embryo-axes-derived calli were capable of producing multiple shoots. However, regeneration frequency and number of shoots per callus were much higher than mature embryo axes.

Proliferating shoots attained a height of 2–3 cm within 8 weeks of culture (Figure 1 *b*). The regenerated shoots were transferred to previously established rooting medium (MS + 5 µM IBA)[9] for adventitious root induction. The rooted shoots reached up to 62% within 4–5 weeks of culture. Regenerated shoots were transferred to pots containing soil mixed with compost. During the first one week, the potted plantlets were covered with glass



Figures 1 *a–c*. Plant regeneration from callus cultures in *A. marmelos*. ***a***, Six-week-old hard green and compact irregular shaped callus induced on MS + 10 µM 2,4-D + 0.5 µM BA. ***b***, Development of multiple shoots from subcultured immature-embryo-axis-derived callus on MS + 5 µM kinetin + 0.5 µM IAA six weeks in culture. ***c***, A potted plant after four months of transfer.

Table 1. Morphogenetic responses of 6-week-old mature- and immature-axes-derived calli of *A. marmelos* to various concentrations of cytokinins and auxins. Each treatment consisted of two replications with 12–18 explants per replication

		Immature embryo axes		Mature embryo axes	
Concentration of cytokinin (μM)	Concentration of auxin (μM)	Percentage calli with shoots	Mean number of shoots per callus ± SE	Percentage calli with shoots	Mean number of shoots per callus ± SE
BA 2.5	IAA 0.5	40.1	38.7 ± 2,5	35.0	17.2 ± 1.8
5.0	0.5	50.3	48.7 ± 1.9	40.2	20.3 ± 2.0
10.0	0.5	54.0	41.2 ± 2.9	42.8	28.2 ± 2.2
25.0	0.5	50.8	29.6 ± 2.4	38.5	21.6 ± 2.3
BA 2.5	NAA 0.5	36.4	15.2 ± 1.2	25.6	8.2 ± 0.9
5.0	0.5	40.3	24.2 ± 2.7	30.2	14.2 ± 2.4
10.0	0.5	44.0	20.8 ± 2.6	30.8	13.6 ± 2.0
25.0	0.5	40.9	18.6 ± 1.8	28.9	10.2 ± 2.1
Kinetin 2.5	IAA 0.5	54.2	24.6 ± 2.0	41.2	19.6 ± 1.9
5.0	0.5	94.2	56.2 ± 2.1	49.6	26.2 ± 1.8
10.0	0.5	78.5	44.8 ± 2.4	52.5	30.4 ± 1.7
25.0	0.5	60.4	34.6 ± 2.1	44.2	20.8 ± 3.0
Kinetin 2.5	NAA 0.5	50.4	18.6 ± 1.4	35.2	8.2 ± 1.4
5.0	0.5	58.9	30.9 ± 1.8	40.1	14.2 ± 2.6
10.0	0.5	64.0	30.3 ± 3.0	36.0	14.8 ± 2.1
25.0	0.5	64.2	28.0 ± 2.9	37.2	11.4 ± 1.9

beakers to provide high humidity. Transplantation success was 55%. Plantlets were subsequently transferred to larger pots (Figure 1 *c*) and gradually acclimatized to outdoor conditions.

The results presented here indicate that *in vitro* regeneration of complete plantlets is possible using callus cultures from embryo axes of *A. marmelos*. The physiological state of the explant plays an important role in the regeneration of plants. Both immature and mature embryo-axes-derived calli were capable of producing plants. However, while the regeneration frequency of immature-embryo-derived calli was 34.4–94.2%, the morphogenetic capacity of the mature-embryo-axes-derived calli was lower (25.6–52.5%). The immature embryos are known to be a good source for initiating callus which possesses high regenerative capacity, as has been shown in peach[11], oak and linden[12]. The procedure described here may be suited to use in transformation studies and for mass propagation of *A. marmelos* if the frequency of off-types is low.

1. Hooker, J. D., *The Flora of British India*, Reeve and Co., England, 1875, vol. 1, pp. 158–159.
2. Biswas, K. and Ghosh, A., *Bharatio Bonoushadhi*, Calcutta University Press, Calcutta, 1973, vol. 1, pp. 187–190.
3. Shorma, D., *Shyamolee Nisharga*, Bangla Academy, Dhaka, 1980, pp. 141–144.
4. Arya, H. C., Ramwat, K. G. and Suthar, K. C., *Indian Bot. Soc.*, 1981, **60**, 134–137.
5. Singh, U. R., Panday, I. C. and Prasad, R. S., *Punjab Hort. J.*, 1976, **16**, 57–59.
6. Islam, R., Karim, M. R., Hossain, M. and Joarder, O. I., *Plant Tissue Cult.*, 1992, **2**, 55–59.
7. Islam, R., Hossain, M., Joarder, O. I. and Karim, M. R., *J. Hort. Sci.*, 1993, **68**, 495–498.
8. Hossain, M., Karim, M. R., Islam, R. and Joarder, O. I., *Plant Cell Tissue Organ Cult.*, 1993, **34**, 199–203.
9. Hossain, M., Islam, R., Karim, M. R., Joarder, O. I. and Biswas, B. K., *Scientia Hort.*, 1994, **57**, 315–321.
10. Murashige, T. and Skoog, F., *Physiol. Plant.*, 1962, **15**, 473–497.
11. Hammerschlag, F. A., Bauchan, G. and Scroza, R., *Theor. Appl. Genet.*, 1985, **70**, 248–251.
12. Chalupa, V., *Plant Cell Repr.*, 1990, **9**, 398–401.

R. ISLAM
M. HOSSAIN
M. R. KARIM
O. I. JOARDER

Department of Botany,
University of Rajshahi,
Rajshahi 6205, Bangladesh.

In vitro axillary bud break and multiple shoot production in *Acacia auriculiformis* by tissue culture technique

Acacia auriculiformis A Cunn. Ex. Benth is a species native to Australia, Indonesia and Papua New Guinea. It has been an important exotic in many tropical countries for more than half a century. The species is currently planted for paper pulp industry[1], fuel wood, erosion control, shade and ornamental purposes. It is a leguminous, nitrogen-fixing tree of the subfamily Mimosoideae. Mass multiplication of elite and rare trees is of paramount importance[2]. Though *Acacia auriculiformis* is propagated by seeds, poor seed germination, around 20–25% has been recorded[3]. So, research is needed on large-scale multiplication of elite clones. The tissue culture technique can be used as a promising tool in overcoming this problem. Much less work has been carried out in this regard in Acacia species. A few workers have reported plantlet formation at seedling level through cotyledonous callus, shoot tip callus and axillary buds[4–6]. But no work has been reported on plant regeneration in mature Acacia trees. The present study reports on *in vitro* development of nodal shoots from axillary buds and multiplication of these nodal shoots into plantlets from 20-year-old Acacia trees.

Small tender branches (10–15 cm length, with 6–7 nodal segments) with

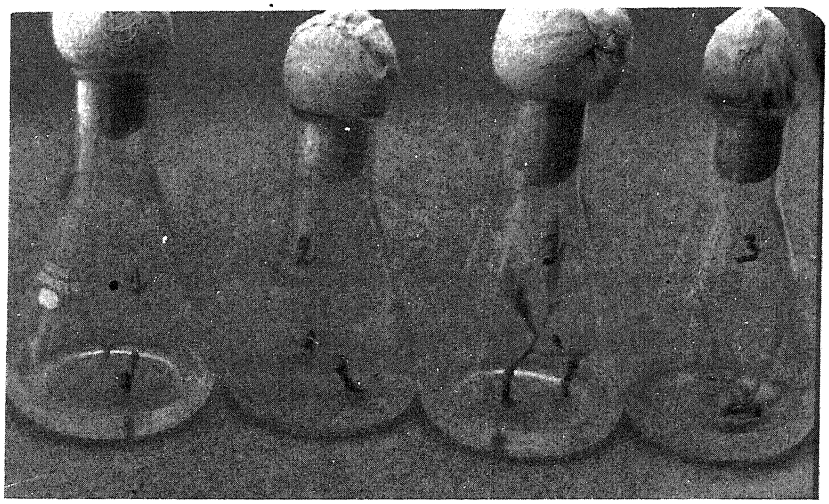


Figure 1. Development of shoots from nodal segments.

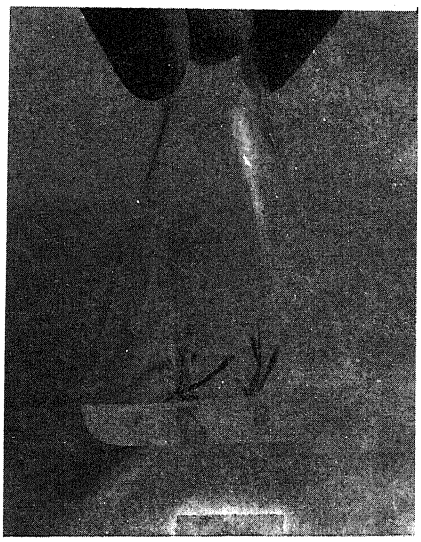


Figure 2. Development of multiple shoots from nodal segment shoot.

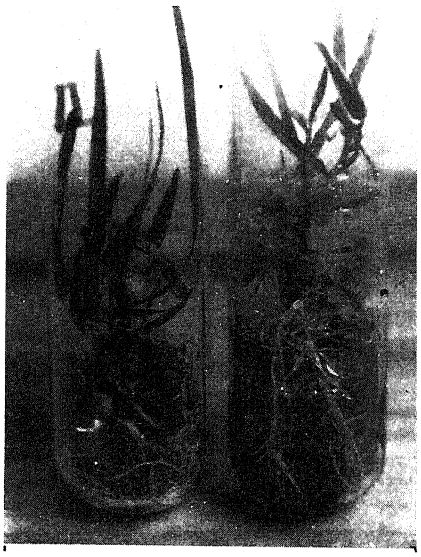


Figure 3. Development of roots in *A. auriculiformis*.

buds of *Acacia auriculiformis* were cut from the mother plant and washed thoroughly under running tap water for 1–2 h. Then they were treated with Bavistin (W/P) 0.5% for 1 h and washed thoroughly in single-distilled water 2–3 times. These branches were surface-sterilized aseptically with 0.1% mercuric chloride for 8–10 min and then thoroughly washed with sterilized water 4–5 times. The leaves were removed and the branches cut into small nodal segments. These segments were placed on modified Murashige and Skoog medium[7] with 0.2% gelrite/0.8% agar supplemented with BAP (1 and 2 μM per litre) and NAA (1 and 2 μM per litre), and 2% sucrose; the medium was adjusted to 5.8 pH and autoclaved for 15 min at 15 lb/in^2 and a temperature of 120°C. Then these flasks were incubated under 16 h light and 8 h dark period at a temperature of 26 ± 3°C. After one month the axillary buds sprouted and developed into shootlets (Figure 1). These shootlets were excised from the main explant and transferred to the multiplication medium containing MS medium with high phosphate (NaH_2PO_4 340 mg/l), supplemented with BAP (5.0 and 7.5 μM/l) and NAA (1 and 2 μM/l), 2% sucrose and 0.2% gelrite. Twenty days after inoculation, 5–7 small shoots per culture arose from the base (Figure 2). Research is in progress to achieve more multiple shoots by altering the cytokinin and auxin ratio. Rooting was achieved in MS medium with charcoal (0.1%) and MS medium with 0.5 mg IBA and charcoal at 0.1% (Figure 3).

1. National Academy of Science, *Tropical Legumes Resources for the Future*, NAS, Washington, DC, 1984, pp. 165–171.
2. Bajaj, Y. P. S., in *Biotechnology in Agriculture and Forestry* (ed. Bajaj, Y. P. S.), Springer, Berlin, 1986, vol. I, pp. 1–23.
3. Sur, K., Lahari, A. K. and Basu, R. N., *Indian Agricult.*, 1987, **31**, 115–122.
4. Das, P. K., Chakravarthi, V. and Mally, S., *Indian J. Forestry*, 1993, **16**, 189–192.
5. Skolemen, R. G. and Mapes, M. O., *Indian J. Herd.*, 1976, **67**, 114–115.
6. Mittal, A., Agarwal, R. and Gupta, S., *Plant Cell Tiss. Organ Cult.*, 1989, **19**, 65–70.
7. Murashige, T. and Skoog, F., *Physiol. Plant.*, 1962, **5**, 473–497.

P. CHANDRASEKHARA REDDY
VEERANAGOUDA PATIL
T. G. PRASAD
K. PADMA
M. UDAYAKUMAR

Department of Crop Physiology, University of Agricultural Sciences, GKVK, Bangalore 560 065, India.

Effect of foetal exposure to low-dose X-rays on the postnatal growth of mouse

Even though the late organogenesis period in mammals is considered to be the most sensitive one to radiation-induced developmental anomalies, a recent study from our laboratory has shown that the early foetal period, day 14.5 post-coitus (p.c.), is also susceptible to γ-ray-induced foetal growth retardation and microcephaly in mouse

Table 1. Observations on postnatal development of mice exposed to 0.05 Gy of X-rays at day 14.5 post-coitus

Observations	Treatment	Age of offspring (week)						
		0	1	2	3	4	5	6
Number of offsprints	C	122	121	120	119	117	115	114
	E	98	96	95	94	92	88	87
Growth-retarded offsprings(%)	C	1.64(2)	1.65(2)	2.50(3)	3.36(4)	3.42(4)	3.48(4)	3.51(4)
	E	4.08(4)	4.17(4)	7.37(7)	10.64(10)*	11.96(11)*	11.36(10)	8.04(7)
Body weight (mean ± SE, g)	C	1.67 ± 0.014	4.76 ± 0.054	7.27 ± 0.081	11.59 ± 0.159	17.57 ± 0.367	24.43 ± 0.273	28.05 ± 0.303
	E	1.64 ± 0.018	4.62 ± 0.059	7.09 ± 0.125	10.71 ±0.215	15.68 ± 0.486	22.23 ± 0.489[c]	27.15 ± 0.446
Body weight (mean ± SE, mm)	C	32.1 ± 0.115	46.78 ± 0.235	54.91 ± 0.227	68.06 ± 0.366	79.65 ± 0.447	87.83 ± 0.366	91.64 ± 0.417
	E	31.88 ± 0.135	46.19 ± 0.317	54.79 ± 0.417	64.68 ± 0.670	71.62 ± 0.789[c]	83.95 ± 0.949[b]24.	89.89 ± 0.814
Head length (mean ±SE, mm	C	8.73 ± 0.056	13.83 ± 0.091	15.86 ± 0.103	20.85 ± 0.146	22.38 ± 0.160	12 ± 0.115	24.70 ± 0.125
	E	8.58 ± 0.061	13.60 ± 0.119	16.19 ± 0.146	20.59 ± 0.122	21.13 ± 0.185[c]18.8	23.58 ± 0.176[a]	24.36 ± 0.134
Head width (mean ± SE,mm)	C	7.99 ± 0.075	12.31 ± 0.075	15.58 ± 0.133	18.03 ± 0.168	7 ± 0.214	19.98 ± 0.162	20.96 ± 0.167
	E	7.99 ± 0.096	12.11 ± 0.146	14.86 ± 0.239[b]	16.83 ± 0.204[c]	17.71 ± 0.257[c]	19.23 ± 0.243[a]	20.39 ± 0.223
Tail length (mean ± SE, mm)	C	13.25 ± 0.140	26.13 ± 0.137	42.37 ± 0.246	57.45 ± 0.371	70.26 ± 0.429	79.63 ± 0.417	84.21 ± 0.460
	E	13.06 ± 0.136	25.57 ± 0.195	40.72 ± 0.445[a]	54.93 ± 0.638[b]	68.26 ± 0.635[a]	74.08 ± 0.806[c]	80.19 ± 0.718[c]

Note: Figures in parentheses are the actual numbers.
[†]C: Sham-irradiated animals, number of mothers 15.
E: Exposed to 0.05 Gy X-rays, number of mothers 12.

Difference from respective controls (C): [a]$p < 0.05$, [b]$p < 0.01$, [c]$p < 0.05$, (Mann–Whitney test) *$p < 0.05$ (Fisher's exact test)

(unpublished observation). Therefore a study was done to find out the effects of low-dose X-rays at day 14.5 p.c. on the postnatal growth of mouse.

Virgin female and male Swiss mice (8–10 weeks old) from an inbred colony, maintained under conditions of controlled temperature (22 ± 3°C) and light (10 h light, 14 h dark), were mated overnight. The day the vaginal plug was observed was considered as day 0 of pregnancy. Pregnant females were exposed to 0.05 Gy of X-rays on day 14.5 of gestation, as described earlier[1,2]. One group of animals was sham-irradiated to serve as control. The experimental and control animals were left to parturate. The offsprings were observed for changes in body weight, body length, head length, head width, tail length and post-natal mortality from day 0 (the day of birth) up to 6 weeks post-partum (p.p.). Sex ratio was determined at birth. The data were evaluated by the non-parametric Mann–Whitney *U*-test and Fisher's exact test, using GraphPad In-Stat software; p values < 0.05 were considered statistically significant.

A total of 220 offsprings from 27 mothers were examined. The litter size, sex ratio and postnatal mortality were not affected by irradiation. The number of growth-retarded offsprings was significantly ($p < 0.05$) higher than control at 3 and 4 weeks, but the difference became nonsignificant by 5 weeks (Table 1).

Irradiation did not affect the birth weight of the offsprings. The body weight showed significantly ($p < 0.01$, $p < 0.001$) lower than control values from 3 to 5 weeks (Table 1). A similar response was observed for body length also. The head length showed significantly ($p < 0.001$, $p < 0.05$) lower values at 4 and 5 weeks of age. However, reduction in head width was apparent from 2 to 5 weeks. All the parameters recovered by 6 weeks. The tail length was significantly ($p < 0.05$) lower than control at 2 weeks, with further reduction at later postnatal ages, with no recovery (Table 1).

Based on the early data, it was concluded that the killing of foetuses requires high doses of radiation[3]. However, more recent *in vitro* and *in vivo* studies have shown that relatively small doses (5–10 cGy) can cause embryonic and foetal mortality[4]. The present data did not show any significant increase in postnatal mortality.

A decrease in weight of rats and mice has been reported after exposure to higher doses of radiation[5,6]. Studies with very low doses are few. Hande *et al.*[2] observed that 5 cGy of X-rays given at the major organogenesis period induced a significant reduction in postnatal weight. The present study shows that such an effect can be induced by low-dose X-rays when exposed at the early foetal period also, but this effect is temporary and complete recovery is effected by the time the animals reach adulthood. There are no comparable published reports on postnatal body length and head size measurements, though our own study has shown a reduction in the head size of 18-day-old foetuses after exposure to ~1 cGy X-rays during the late organogenesis period[1]. Diminished head size has been reported to be a readily detectable effect in prenatally exposed Japanese A-bomb victims[7].

Tail growth seems to be more sensitive to exposure at this stage than the other parameters. There are no comparable studies on the effect of low-dose *in utero* exposures during early foetal period on postnatal tail measurements. However, stunting of skeletal growth has been reported by several investigators after high doses during 14–18 days p.c. in rats and mice[8,9].

Recently, Uma Devi *et al.*[10] have reported that doses below 25 cGy of γ rays produced no marked foetal mortality or growth retardation when exposed on day 11.5 p.c., which is a more radiosensitive stage than day 14.5 p.c. in mouse development. However, the present results suggest that in the case of X-rays even doses as low as 5 cGy on day 14.5 p.c. can lead to growth retardation. Even lower doses (~1 cGy) of X-rays during the organogenesis period have been found to produce growth retardation in mouse foetus[1, 11]. These results strengthen our earlier assumption that at low doses X-rays may be more effective in influencing mouse development than γ-radiation[10].

1. Uma Devi, P. and Hande, M. P., *Experientia*, 1990, **46**, 511–513.
2. Hande, M. P., Uma Devi, P. and Jagetia, G .C., *J. Radiat. Res.*, 1990, **31**, 354–360.
3. United Nations Scientific Committee on the Effects of Atomic Radiation, United Nations, New York, 1977.
4. United Nations Scientific Committee on the Effects of Atomic Radiation, United Nations, New York, 1986.
5. Rugh, R., Duhamel, L. and Skaredoff, L., *Soc. Exp. Biol. Med.*, 1966, **121**, 714–718.
6. Sikov, M. R., Resta, C. F. and Lofotrom, J. E., *Radiat. Res.*, 1969, **40**, 133–148.
7. Miller, R. W. and Blot, W. J., *Lancet*, 1972, **2**, 784–787.
8. Murphree, R. L. and Pace, H. B., *Radiat. Res.*, 1960, **12**, 495–504.
9. Nash, D. J. and Growen, J. W., *Biol. Bull.*, 1962, **122**, 115–136.
10. Uma Devi, P., Baskar, R. and Hande, M. P., *Radiat. Res.*, 1994, **138**, 133–138.
11. Michel, C. and Fritz-Niggli, H., *Fortschr. Röntgenstr.*, 1977, **127**, 276–280.

ACKNOWLEDGEMENTS. X-ray facility was provided by the Radiology Department, Kasturba Hospital, Manipal. Technical assistance by Mrs Sandhya and Ms Geethanjali is thankfully acknowledged.

K. S. BISHT
P. UMA DEVI

Department of Radiobiology,
Kasturba Medical College,
Manipal 576 119, India.

The quest for quantum gravity

Gary K. Au

One of the greatest challenges facing theoretical physics lies in reconciling Einstein's classical theory of gravity – general relativity – with quantum field theory. Although both theories have been experimentally supported in their respective regimes, they seem mutually incompatible. This article summarizes the current status of the superstring approach to the problem, the status of the Ashtekar programme, and addresses the problem of time in quantum gravity. It contains interviews with Abhay Ashtekar, Chris Isham and Edward Witten.

'It is very important that we do not all follow the same fashion ... It's necessary to increase the amount of variety ... and the only way to do this is to implore you few guys to take a risk with your lives that you will not be heard of again, and go off in the wild blue yonder to see if you can figure it out.'

Richard Feynman (1965), Nobel prize in physics award address

The road to quantum gravity (QG) has had a long and winding history. Although we are unlikely to see the day of a direct test of quantum gravity, such a theory is needed to describe the early stages of the Big Bang. Around this time the universe would be about the size of the Planck volume. Quantum gravity effects, ignored in particle physics because of their weakness, would have a major influence on the subsequent evolution of the universe. Also a QG theory is desperately needed to make the standard model a consistent physical framework.

Why are we spending so much time on such a lofty goal? Richard Feynman commented on the aims of science: 'If you expect science to give you all the answers to the wonderful questions about what we are, where we are going, and what the meaning of the universe is, then I think you could easily become disillusioned and look for a mystic answer to these problems.'

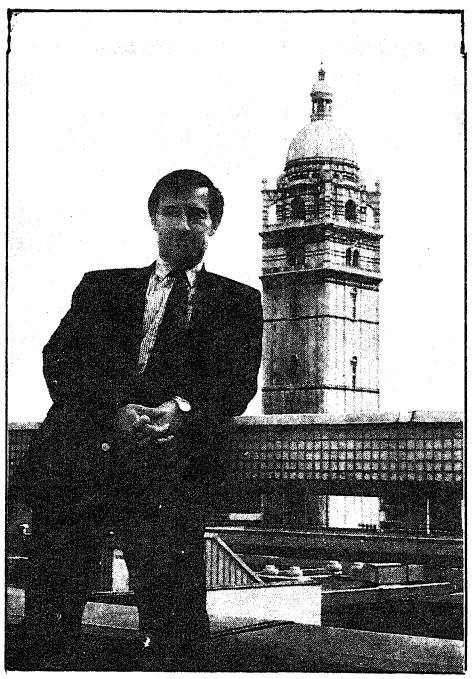
'The way I think of what we're doing is, we're exploring – we're trying to find out as much as we can about the world. People say to me, "Are you looking for the ultimate laws of physics?" No, I'm not. If it turns out there is a simple, ultimate law which explains everything, so be it; that would be very nice to discover. If it turns out it's like an onion, with millions of layers, and we're sick and tired of looking at the layers, then that's the way it is. But whatever way it comes out, it's nature, and she's going to come out the way she is! Therefore, when we go to investigate it we shouldn't predecide what it is we're going to find, except to try and find out more.' [*No Ordinary Genius – The Illustrated Richard Feynman*, Christopher Sykes, W. W. Norton, New York, 1994, p. 251]

Gary K. Au is in the Research Centre for High Energy Physics, School of Physics, The University of Melbourne, Parkville 3052, Victoria, Australia.

The problem of time[1–5]

A key question which faces any approach to QG is the 'problem of time': how does a time variable, with its special properties in relation to the probability interpretation of quantum mechanics (QM), emerge? A related topic which has been attracting increasing attention, especially in the context of 'quantum cosmology' is the 'consistent-histories' approach to quantum theory. One of the most exciting features of the programme is the possibility of changing quantum theory itself in the context of the physics of the early Universe and the way in which it came into being.

The basic problem lies in the incompatible way in which the concept of time is treated in QM and general relativity (GR). One of the main researchers studying this deep problem is Chris Isham at Imperial College, London.



Chris Isham

In QM time evolution is described by the Schrödinger equation. However, the time parameter that comes into the equation is a background parameter. This is unlike other quantities which can be represented by Hermitian operators. So from the very beginning QM attaches a special significance to the idea of time. It is like the steady ticking of an external clock hand in the background, which we have no control over. Can we physically measure this quantity? Is it possible to reformulate QM so that the idea of a background time does not appear? According to Chris: 'It is possible to do this in a formal sort of way but it does not really get you far. The basic fact is that conventional quantum theory presupposes an external time whose ontological (see note 1) status is the same as it possesses in classical Newtonian physics. In relativistic quantum theories this is generalised to become a background Minkowskian spacetime – so it is the causal structure that is fixed. There are some intriguing mathematical theorems (going back to Pauli I believe) showing that you cannot construct quantum clock that would exactly measure this background time. In that sense, some people might want to consider "time" as part of the external classical realm that Bohr postulated was necessary to interpret the equations of quantum theory. I have always suspected that Bohr himself would have regarded the subject of "quantum gravity" as a non starter!'

Why does time deserve a special status in QM? Well, if you could raise it to the status of an operator T then its conjugate variable would be the energy H. These would need to satisfy the commutation relations $[T,H] = i\hbar$. Mathematically, it can be shown that it is impossible to have a pair of Hermitian operators satisfying these conditions, in which one of them has only positive eigenvalues. This relates back to Chris' remark about the problem of quantum clocks. In other words, we have the problem of a system with negative energy.

In QM, measurements of observables are taken at fixed times, a complete set of commuting operators (defining the maximum information you can obtain from a system) is required at fixed times, and the scalar product on the Hilbert space of states has to be conserved under time evolution of the Schrödinger equation. Time plays a foundational role in the technical and conceptual aspects of quantum theory. However, it is an idea grounded in the world of continuum mathematics. If it breaks down at the Planck scale, where some believe a discrete structure will emerge, what are we left with? Indeed, how much of QM will hold?

According to John Wheeler: 'The word "time" was not handed down from heaven as a gift from on high; the idea of time is a word invented by man, and if it has puzzlements connected with it, whose fault is it? It is our fault for having invented and used the word (see note 2). Taking this into mind, could we redefine time so that it reduces to the familiar clock ticking concept in some classical limit?'

Chris Isham is very sympathetic to this view: 'I would bet strongly that whatever the "final" theory of quantum gravity will be, we shall see our standard notion of time emerging only in some semiclassical sense.'

The role of time in relativity is very different to its role in QM. Relativity treats time at par with the other three spatial variables. Indeed, time and space are unified into a four dimensional spacetime manifold.

On this matter Chris explains: 'The notion of a single, four-dimensional, "spacetime" (rather than separated three-dimensional space plus one-dimensional time) first appears in special relativity (SR) because of Einstein's realization that any absolute split of 3+1 is observer-dependent. General relativity works with a curved version of the flat space–time of special relativity.'

'In both special and general relativity a "moment of time" corresponds to a single space-like hypersurface in the space–time. Time itself appears as the parameter that labels the elements of a one-parameter foliation (see note 3) of space-time by such surfaces. In the case of special relativity you restrict your attention to the case where the space-like surfaces are (i) hyperplanes, which are (ii) mapped into each other by actions of the Poincaré group. The main challenge in quantum theory is to show that all such admissible families give the same physical answers.'

'The situation in general relativity is more complex. One says that a hypersurface is space-like if the vectors tangent to each point of the surface are space-like. The idea here is that the Lorentzian metric on spacetime induces on each tangent space a copy of the Minkowski metric of special relativity, and it is with respect to this latter metric that the tangent vector has to be space-like. Put slightly less rigourously, a hypersurface is space-like if any "infinitesimal" transformation in the surface always points in a space-like direction.' (Figure 1)

The crucial point here is that the problems encountered when incorporating SR into QM can be overcome. When we try to add GR to the QM picture, the intricacies are much more difficult to handle.

With a *curved* spacetime equipped with a Lorentzian metric tensor, how many ways are there to foliate spacetime as a one parameter family of space-like surfaces? Chris continues: 'In some spacetimes there is a topological obstruction to performing *any* global foliations. But if such foliations are possible then there will be an uncountable number of them, each corresponding (in general relativity) to an allowed definition of time.' In other words, unlike in SR, we cannot define a unique time according to which we evolve a system. This is the main difference between the concept of time in special and in (classical) general relativity. It is an important aspect to the problem of time in QG. On the one hand, we have a Hamiltonian formulation of QM for privileged SR space-like surfaces. On the other we have general covariance in GR, which states that no one set of

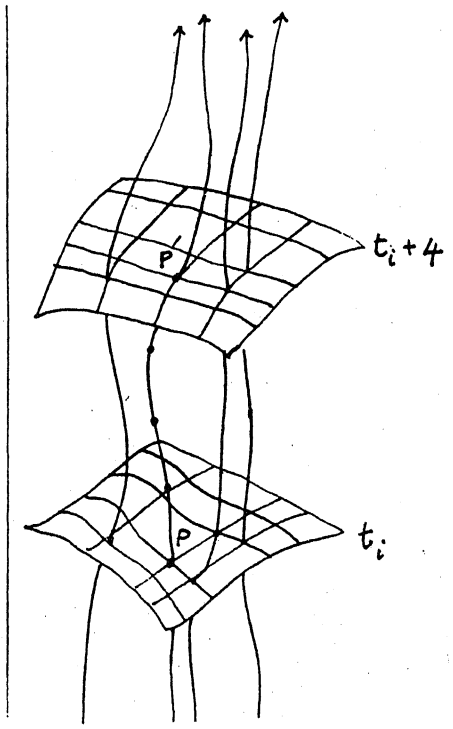


Figure 1. Foliating spacetime into a one-parameter family of space-like hypersurfaces: In Newtonian mechanics, a given moment of time has the same meaning for all observers. Time is absolute. In SR, this notion is somewhat ambiguous because of the nonuniversality of simultaneity. However, once an inertial frame is specified, the notion becomes more precise. Unless spacetime is flat, there are no global inertial frames in GR. So 'a moment of time' in GR is ambiguous. However, a more general concept replaces it; the notion of 'a given space-like hypersurface'. It is possible to 'slice up' or foliate spacetime into a one-parameter family of 3D space-like surfaces, filling out all of spacetime. At each event on a surface, there is a local Lorentz frame whose surface of simultaneity coincides locally with the hypersurface. Lorentz frames at various events do not mesh to form a global inertial frame, but their simultaneity surfaces mesh to form the actual space-like hypersurface. This is a picture of two space-like hypersurfaces through events P and P' which occur at different times. World lines of observers cross these surfaces orthogonally, and are labelled by (x_1, x_2, x_3), $x_0 = t_i + \tau$, where τ is the proper time measured along the worldline. Note that the third coordinate x_3 is suppressed in the diagram. Arbitrary coordinate grids defining x_1 and x_2 are drawn on the hypersurfaces. At the point of crossing a hypersurface, the proper time (defined by an observer's wristwatch) reads the same as other observers' watches. Such a surface defines a notion of 'simultaneity'. Another surface of simultaneity through the event P may have a different slope or curvature (defining different local Lorentz frame). This second surface is defined by another set of observers with their own wristwatches, and their own proper time readings. It is also possible to push the hypersurface ahead in time at different rates in different places, so long as the hypersurface stays spacelike. This idea of a 'many-fingered' time is GR's dramatic new contribution to our understanding of time. See, for example, *Gravitation*, C. Misner, K. Thorne and J. Wheeler, W.H. Freeman and Co., 1973.

space-like surfaces has privilege as a reference system over others.

When we get to Planck scale physics, spacetime geometry could be subject to Heisenberg's uncertainty principle, varying about quantum-mechanically. In other words, the GR metric tensor $g_{\mu\nu}$ could have fluctuating components! How could one define a light cone and hence a space-like, light-like or time-like separation under such circumstances? Without such a definition, the idea of a space-like surface seems invalid. Perhaps the idea of a spacetime (a continuous manifold of points) does not have physical meaning in the QG regime, where the light cone could be smeared out.

To this Chris says: 'You cannot define a light cone, etc., except in some "background" sense, and this perhaps is the essence of the problem of time in QG: there is no fixed microcausal structure which can be used to construct a relativistic quantum theory in any of the standard ways.'

'However, this does not rule out a priori a continuous manifold: the ideas of causal structure and continuous manifold are in no way synonymous. You can easily have the latter without the former (although I myself am very sympathetic to the idea that the idea of a "continuum spacetime" is not something that will carry across to the "final" QG theory).'

One possibility at the Planck scale is for topology changes in spacetime geometry. Fixing a background topology and differential structure to that of Minkowski space, as is done in quantum field theory, seems to presuppose a lot at the Planck scale.

If we stick with a continuum formulation of QM based on complex numbers, could we be missing out on revealing some sort of a discrete structure at the Planck scale? One can see how a discrete theory could reduce to a continuum one in the large scale limit, but to shed light on a discrete theory while working from the perspective of a continuum one seems difficult to achieve.

Chris has felt this for a long time. One of the many reasons why he is so interested in the new decoherent-histories approach to quantum theory, and is developing his own quantum-logical version, is that the role of the continuum is isolated in a particularly efficient way. In his scheme it is just the space in which the decoherence functionals take their values. He says: 'It is sometimes suggested that near the Planck length a more "combinatorial" approach to physics may be appropriate, and I have some sympathy with this view. I think it would be much easier to think about such schemes within this new approach to quantum theory than in the old Hilbert space one.'

Is superstring theory able to handle this fundamental problem of time in its present perturbative formulation?

Chris replies: 'No. Although it is essential to emphasize that the "problem of time" is not a single problem (or even collection of problems) that applies willy-nilly to any attempt to construct a quantum theory of gravity. It is very approach-dependent and looks very different

in different approaches to quantum gravity. For example, if space and time themselves do "emerge" in some nonperturbative way from superstring theory, then questions concerning the nature of time will look very different from those that arise naturally in, for example, the Wheeler–DeWitt approach to canonical quantum gravity. That is one of the many reasons why it is important to keep trying to find a new nonperturbative way of looking at superstring theory.'

The universe is modelled as a closed system. In its present formulation QM is inadequate for handling such models. After all, how does one define a time external to the universe? How successful have attempts been to define an internal time using, say, matter fields?

'They have been getting better. In a recent paper (see note 4) Karel Kuchar (who is the world expert on this type of thing) has come extremely close. Just how close depends precisely on what one decides one is trying to achieve. From one perspective one could say that he has succeeded in a technical sense, although he would be the first one to agree that his particular definition of time is not too compelling in any fundamental physical sense.'

Karel uses dust particles (see note 5) to fix points in space and identify moments in time. The dust acts like a reference fluid analogous to the old aether idea but with the significant difference that the aether was usually regarded as fixed, i.e. part of the background structure, whereas these dust particles do obey equations of motion. Hence, in a sense they are dynamic, as is the geometry of space–time.

One of the problems of working in QG is the lack of empirical data. One clue comes from the 1992 COBE results. These showed variations in the blackbody microwave background coming from different directions of the sky. These date back to 300,000 years after the Big Bang and are consistent with the inflationary scenario. The observations provide evidence for the idea of quantum fluctuations as the origin of the galaxies. Could they have any implication for the QG regime of 10^{-43} s after the Big Bang?

'Stephen Hawking has claimed that they support the ideas on the very early universe put forward by himself and Jim Hartle. However, the COBE data are currently being cited by a number of people in support of a variety of different views, and it is not clear yet what the outcome will be.'

Consistent histories

An approach which attempts to remove reference to external time in QM is that of consistent histories, pioneered by Griffiths, Omnès, Gell-Mann and Hartle. It generalizes QM via a sum over spacetime histories formulation. The motivation is to describe the QM of closed systems in the context of quantum cosmology. It hopes to shed light on the initial condition of the universe. The basic idea in this scheme is a QM history. The scheme, therefore, de-emphasizes the notion of a QM event at a single moment in time. It makes no reference to external observers, classical measurement apparatus, or wave function collapse (reduction of the state vector). However, it is possible to recover Bohr's formulation of QM under suitable conditions. Chris Isham is working on a version of this idea.

'Most work on this approach has been applied to standard quantum theory, not quantum gravity. To bring in quantum gravity, Hartle has discussed the possibility of using solutions of the Wheeler–DeWitt equation in the context of Euclidean path integrals. I myself am working on this general problem but from a completely different tack. I and my colleagues have been studying consistent histories using a new type of "quantum logic" of propositions about generalized spacetime structures. But the general idea in this formalism still holds: i.e. properties can only be ascribed to a "history" only when it is part of a consistent set. Typically, this requires deliberately loosing information by coarse-graining. One's expectation is that "time" is something of this sort: i.e. it is really only part of a semiclassical world that will "fade away" if one probes too closely.'

The scheme's main goal is to assign probabilities to families of histories in a closed system. A history is defined to be a sequence of QM events at a succession of times. QM events are tested for using projection operators which satisfy the properties of exhaustivity and mutual exclusiveness. Two types of projections are defined: fine-grained and coarse-grained ones.

The main problem in the formalism is interference. Interference between different histories forbids assignment of probabilities since probabilities have to satisfy certain axioms to be meaningful. They must be nonnegative, normalized, and they must sum to unity. Complex probability amplitudes can be assigned to histories, but a probability is defined as the modulus squared of an amplitude. This operation introduces cross terms which violate the probability axioms.

To handle this problem 'consistency conditions' are introduced. These determine sets of histories which negligibly interfere, and to which probabilities may be assigned. These conditions can be combined into a decoherence functional, which is the main calculational tool of the QM of history. It must satisfy the properties of hermiticity, positivity, normalization and the superposition principle. The functional is computed using standard quantum mechanics. Griffiths, Omnès, Hartle and Gell-Mann's contribution is the new probability interpretation. Here the notion of 'superposition' does not mean quite the same as it does in the standard way of superposing wave-functions. It is more a question that the decoherence functional must be additive with respect to a disjoint sum of histories.

In order to decouple interfering histories by coarse graining, we require the idea of decoherence. Consistent (decoherent) histories are then sets which obey the consistency conditions. Consistency is a property of families of histories, of which there may be many.

But what is the mechanism of decoherence? It has been postulated (see note 6) that the environment decoheres by acting like a sink, by draining off coherence – similar to a piece of blotting paper soaking up excess ink. But if the system is a closed QM universe, how do you define an environment? What other methods could there be to decohere closed systems?

To this problem Chris remarks: 'The mechanism for decoherence in a truly closed universe is, to my mind, still rather problematic. There have been attempts, for example, to identify part of the gravitational-field complex as the environment, and part as "really quantum" freedom in this environment. But I am doubtful whether this is really a universal notion. The examples I have seen have all been for very simplified models.'

In a closed QM universe, would it be possible to use black holes as decoherence devices? When matter goes into a black hole, the only properties of the hole which are observable via its distant electromagnetic and gravity fields are mass, charge and angular momentum. All other information is lost. Would this loss include 'interference' information between histories?

Unfortunately, Chris says: 'One could do that when studying quantum theory in a fixed background, since then one could consider a background containing black holes. However, this is not too meaningful a priori in a full theory of quantum gravity itself since the black holes will also have to be quantized in some sense.'

Jim Hartle made the important observation that one can talk about 'generalized histories' in QG in which the word 'history' need not presuppose a conventional type of time parameter. According to Chris: 'The basic entity in the scheme is a "possible universe". A simple, but informative example would be a Lorentzian geometry that is not globally hyperbolic. Such a spacetime does not admit a global foliation by space-like surfaces, and hence there is no global time parameter. Nevertheless, it is a viable classical possibility since time still makes sense locally (i.e. the local proper time along a world line). This was one of the original motivating examples suggested by Hartle.'

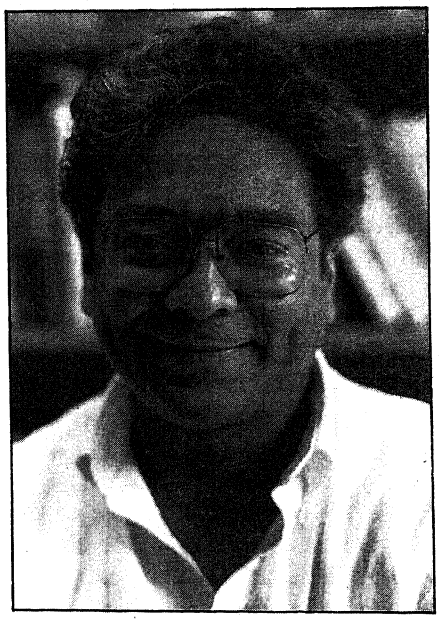
'However, my own interest in this scheme is when the basic entity is something more basic that does not necessarily include any manifest time-like features at all. This would be in line with the general idea that time appears only in some semiclassical limit. Of course, this leaves quite open what types of object one would actually use as "possible universes" since this depends entirely on the actual theory that is to be developed (the decoherent-histories approach is not a *theory per se*, it is more a *theoretical framework* in which new types of theory can be developed). For example, I am currently thinking about the possibility of constructing such a theory in which the basic entities are simply point-set topologies.'

'In theories of this generalized type, Hilbert spaces do not appear as basic mathematical structures in the same way as they do in normal quantum theory; indeed, in the scheme that I am developing, the basic mathematical structure is much closer to that of quantum logic (i.e. an algebraic approach) than it is to Hilbert space methodology. On the other hand, in any theory in which time could be shown to 'emerge' in some coarse-grained limit I would expect the usual Hilbert space structure to emerge too at the same level. However, as no one has yet written down an explicit example of this type, it is hard to be more specific!'

The Ashtekar programme[6–21]

People have approached the QG problem from many different directions, some more successful than others. Most attempts suffer from seemingly insurmountable problems. A short list of avenues includes higher-derivative Lagrangians, twisters, induced gravity, Kaluza-Klein theories, Euclidean quantum gravity, covariant perturbation theory, discrete gravity, nonlinear quantum mechanics, spin networks, asymptotic quantization, quantum cosmology, the decoherent histories just covered, quantum field theory in curved spacetime, superstrings, and canonical gravity. The Ashtekar programme belongs to the last category.

Abhay Ashtekar is at the Centre for Gravitational Physics and Geometry, Pennsylvania. Although his programme involves esoteric mathematics, some of its key features can be described qualitatively.



Abhay Ashtekar

An obvious question that people might ask about QG research is: Why are we spending so much time and effort trying to obtain a theory which is most probably untestable, and which will not have any practical application? Abhay doesn't quite agree with the assumptions here. He says: 'We are seeking a *physical* theory and, therefore, it will definitely have testable predictions. The initial tests will be probably indirect. For example, we have this great puzzle for over 50 years: quantum field theory provides incredibly accurate predictions in particle physics, yet it is, in a sense, only a set of calculational recipes, without a coherent, complete mathematical framework. Again, the recipes work incredibly well and should be taken seriously. But surely, they are incomplete in some essential way. The ultraviolet divergences arise largely because we assume that space–time is a continuum at all scales. If a quantum theory of gravity provides an alternate picture of space–time, it should lead to an alternate way of doing quantum field theories which is coherent and mathematically consistent. So, in a sense, the vast experimental data we have accumulated in particle physics could provide tests of the space-time models that come from quantum gravity.

'Furthermore, it is possible that quantum gravity will also lead to "practical applications". Recall that when Einstein discovered special relativity, most people thought that it was an esoteric theory which would have little impact, if any, on daily life. However, the theory then led to $E = mc^2$, which has had profound ramifications on the entire world order of this century. Quantum gravity will also change our understanding of space and time radically and may, therefore, have some really deep practical implications. I would hope that they would be more peaceful in nature!'

'But of course we do not have the faintest clue of such applications today and, as you indicate, it is beyond today's technology to test the quantum gravity effects *directly*. The primary motivation is really conceptual. Again, there is a similarity with special relativity. Just as the primary motivation for special relativity came from the incompatibility between Newtonian mechanics and Maxwell's electrodynamics, the two pillars of the 19th century physics, the primary motivation for quantum gravity comes from the tension between general relativity and quantum mechanics. There should be a deeper theory which unifies the principles of both in such a way that these two theories emerge in suitable approximations. That is the theory that workers in quantum gravity are trying to construct.'

Abhay's programme was inspired by the failure to find a renormalizable perturbative theory of quantum GR (QGR). What this is, and what would have happened if this failed programme had been successful?

He explains in depth: 'Perturbation theory assumes that answers to physical questions can be obtained by a method of successive approximations, generally expressed in terms of a power series in the coupling constant. We routinely use such approximations already in quantum mechanics. In quantum field theory, however, there is a key difference. Whereas in quantum mechanics, we know from general principles that the exact answer exists and is finite, and perturbative techniques are used only as a computational tool, in realistic quantum field theories in four dimensions, we do not have an assurance that there is an underlying well-defined theory; the perturbation series itself is being used to define the theory. To make matters worse, each term in the perturbative expansion diverges because one is allowing virtual processes of arbitrarily large momenta. In renormalizable theories, the infinities that arise are of a special nature; they can be absorbed into a finite number of parameters associated with the theory such as charges, masses and coupling constants. Once these parameters are renormalized by the appropriate infinite factors, all individual terms in the perturbative expansion become finite. The actual renormalized parameters can be determined by a finite number of experiments. Further experiments then provide tests of the theory. In a nonrenormalizable theory, this cannot be achieved and, therefore, the perturbation expansion has no predictive power.'

'Another way of stating the difference is that, in a renormalizable theory, predictions for phenomena at a given length scale are insensitive to what is happening at much smaller scales. So, for example, since QED is renormalizable, its predictions for processes at, say, a GeV scale are insensitive to all the goings on at the Planck scale (10^{19} GeV). In nonrenormalizable theories, this is not the case; the predictions for one length scale *are* sensitive to what is happening at much smaller, scales. It is thus harder to extract the physical content.'

'If perturbative general relativity had turned out to be renormalizable, on the one hand, life would have been simple. One could have calculated scattering cross-sections in gravity using familiar methods. On the conceptual side, however, at least some people, including me, would have been disappointed. It would have meant that most of physics is insensitive to the "true" small scale structure of space–time and hence it would have been harder to "probe" this structure. I have expressed my unease with the current status of quantum field theory. It is the nonrenormalizability of gravity that forces us to seek the true microstructure of space–time and holds clues for obtaining something which goes beyond (the calculational recipes of) quantum field theory.'

Perturbation theories assume that spacetime is a continuum at all scales. Could this be a reason why they fail for QG, where the short distance behaviour of spacetime might not be a continuum? An analogy has been made that spacetime is a 'foam' of sorts at Planck scales, subject to Heisenberg's uncertainty principle.

'That's right! The key infinity in perturbative treatments is the ultraviolet, one, where processes involve arbitrarily large energies. Surely, when the energy becomes bigger, gravity should become more and more important, and the space–time geometry, more and more nontrivial. Yet, in the perturbative treatments, one uses a fixed, continuum background space–time. Because space–time geometry is a dynamical entity in general relativity, it seems clear that in quantum gravity we cannot *presuppose* what the microstructure of geometry must be; we should let the theory itself tell us. Of course, one might have been lucky and made the right guess at the start. But the failure of perturbative methods tells us that this was not the case with the continuum hypothesis.'

Might quantum GR exist nonperturbatively, i.e. as an exact theory?

'Yes, it might. We do not have any clear evidence to the contrary. Incidentally, there do exist theories in three space–time dimensions which are nonrenormalizable but exactly soluble. Note that I am *not* saying that quantum GR must exist nonperturbatively. However, since there are qualitative reasons to expect that the nonperturbative theory will be very different from the perturbative one, the question is well worth investigating. More importantly, even if the answer turns out to be in the negative, the exercise will have been extremely worthwhile because it has already provided many new mathematical and conceptual tools to construct quantum field theories without a background space–time. Like many others, I firmly believe that, at a fundamental level, the final solution will have to be background-independent; space–time will emerge only as an approximation. And we have very little experience with quantum field theories which have an infinite number of degrees of freedom and which do not use a background space–time.'

Do you see the need for a fundamental revision of the present concepts on which the standard model is based? In other words, are we in need of a conceptual revolution to solve this problem? I once spoke to Bill Unruh, and he did not think QG could be obtained through 'tinkering with mathematics'. He said we needed to feed in new conceptual ideas.

According to Roger Penrose '...if there is to be a final theory, it could only be a scheme of a very different nature. Rather than being a physical theory in the ordinary sense, it would have to be a principle – a mathematical principle whose implementation might itself involve nonmechanical subtlety.' Is this your point of view as well?

'I believe that, at a fundamental level, the continuum picture has to go. This would be a profound change, in terms of both physics and mathematics since the picture is embedded so deeply in the conceptual fabric of physics. So, I agree with Bill that "tinkering" with mathematics alone will not suffice. However, I also feel that new mathematics *will* be needed. (I mean mathematics that was not previously used in physics.) This has happened with most big breakthroughs in physics. Newtonian physics needed calculus, general relativity needed differential geometry and quantum mechanics, Hilbert spaces and operator algebras. Indeed, without access to new mathematics, it would not have been possible to formulate new questions, let alone analyse them! Certainly, what Roger suggests is a possibility.'

Most approaches to the QG problem seem to take QM at face value, and try to modify GR in some way, or take GR as a weak-field limit. This is the case for superstring theory, where GR appears as a low-energy limit. The dimensional nature of the basic Planck units – the Planck length, time, and energy – lends credence to the idea that a QG theory could reproduce GR in regimes well away from the Planck scale. If QM is only an approximate theory itself, which fails to hold at Planck scales, then in the spirit of Roger Penrose's suggestion, QM will also have to undergo a revision. Abhay feels that this is in part why *all* attempts at quantum gravity that are being pursued are likely to be incomplete. 'However', he says, 'it is very difficult to provide concrete models of how QM should be revised. I think the best bet is to push all viable approaches to the end and if cracks develop, they will suggest how we should change QM.'

Loop variables gravity seems like a conservative approach compared to superstring theory. One major difference is that according to your programme, quantum gravity can be obtained without a unification of all forces. In the superstring theory, such a unification is vital for consistency. To be consistent, superstrings also need supersymmetry (a mathematical relationship between bosons and fermions) and higher spacetime dimensions (nine space and one time). The extra six spatial dimensions are supposed to be compactified into a very small region. Do you see these differences as complementary or conflicting issues? 'On the whole, I think the differences are complementary. String theory provides powerful constraints on allowable interactions. On the other hand, it has proved difficult to analyse string theory nonperturbatively, without assuming a background space–time. As I indicated above, our approach provides concepts and tools for a background independent approach. Over the last year or so, more concrete calculations have been proposed which may pull the two approaches closer.'

'I should also point out that we do not have a proof that quantum GR exists nonperturbatively. However, the tools required to analyse this issue finally exist. It is possible that we will find e.g. that supersymmetry or some specific couplings are necessary. This year is an exciting one for our programme because we have begun to analyse such issues.'

I spoke to Peter Bergmann (one of the founders of the geometrodynamical route to QG, and a coworker of Einstein) while I was an undergraduate and he was very skeptical of superstrings. Is it the same with you, considering most of the work in this field has been grounded in perturbation theory?

'I am perhaps more impressed than Peter by the mathematical successes of string theory. I agree that the use of a background space–time is a severe weakness. However, the very recent results due to Ashoke Sen (see note 7) and others on black holes within string theory are providing some glimpses of the possible relation between perturbative results and nontrivial geometries.'

Unlike the other forces, the gravitational force is nonlinear. In contrast, the electrical forces are linear because the total force due to a number of charges is just the sum of the forces due to each charge acting alone. This situation corresponds to the vector law of addition. Now in the case of gravity, we have a nonlinear combination of forces. This is because all forms of energy act as a source of gravity and inertia. This is true for the gravitational potential energy existing between two separated masses. Thus, the total gravitational force involved when two objects interact does not follow the usual linear vector addition law. Is this nonlinearity one of the root causes of all the problems in quantizing gravity?

'This profound nonlinearity is part of the difficulty. But this part is technical. The most important difficulty, in my view, is conceptual: it lies in the fact that gravity is encoded in the very geometry of space–time. To quantize gravity, one has to quantize geometry. One has to learn to do physics and mathematics in the absence of space–time.'

I guess one of the things the readers want to know is how everything reduces down to the picture of an apple falling from a tree. Newton's idea of a 'force' pulling the apple down is intuitive. In GR, this picture was replaced by the concept of the curvature of a four-dimensional spacetime continuum. The apple's mass curves the spacetime, and the Earth's mass curves the spacetime. This curvature tells how the masses should move. The two masses feel the combined curvature, and fall towards each other. In QFT, this 'attractive' picture is replaced by the concept of the exchange of particles called gravitons. The gravitons mediate the gravitational pull. Does the loop variables theory preserve this idea?

'Yes. Although in our approach gravitons do not exist at a fundamental level, in the low-energy approximation one can identify certain loop states with gravitons and then the picture is the one you indicate.'

'Incidentally, the loop variables also have a direct experimental significance. If you move an electron around a closed loop in a gravitational field, its spin undergoes a rotation. The basic loop variable just measures this rotation. Furthermore, if you know the rotation for all closed loops, you know the gravitational field completely.'

One of the problems with gravitons is that they seem to make sense only when the gravitational field of GR is linearized, i.e. split up into a flat space metric, and a perturbation. Gravitons are like ripples in a flat space–time background. Conventional particle theory tries to quantize these undulations, while leaving the background classical. When the background starts curving noticeably itself, what reference is there for these ripples? The problem seems to be that the arena and performers cannot be distinguished anymore.

'Gravitons are indeed ripples on a space–time background. In the perturbative approach, one does quantize these undulations, leaving the background untouched. In our approach, the situation is different. Very roughly speaking, Minkowski space by itself is represented as a quantum loop state – which we call a weave – and certain nearby states – embroidered weaves – can be thought of as gravitons propagating in Minkowski space. However, as I indicated above, the picture works only for low energy gravitons. In this regime, one can distinguish the arena from the performers.'

Geometrodynamics

Abhay's programme takes conventional GR coupled to matter as the leaping point. It is based on the 'canonical' approach to QG, whereby the full four-dimensional spacetime is split up into an observer-dependent 3+1. The first step in any canonical approach is the Hamiltonian formulation of the theory. This was achieved in the early sixties by Dirac, Bergmann, Arnowitt, Deser, Misner and others. In their formulation, the basic variable is the metric on the three-dimensional space; it captures the intrinsic geometry of the hypersurface. The Hamiltonian framework was, therefore, baptized 'geometrodynamics' by Wheeler. Unfortunately, in this framework, the equations of the theory are complicated and, therefore, as far as the full theory is concerned, the quantization programme itself did not really take off.

In geometrodynamics the canonical variables are the intrinsic geometry and the extrinsic curvature of the 3D spatial hypersurface (which is embedded in a 4D Riemannian spacetime). The subject has close links with Gauss' theorem Egregium.

Abhay explains: 'The intrinsic geometry of a three-dimensional hypersurface is the geometry of space that we can "experience directly". It tells us, for example, how to measure lengths of curves lying in the surface. The extrinsic curvature, on the other hand, tells us how the surface is embedded in the four-dimensional space–time. To see the difference, let us go one dimension down and consider a cylinder embedded in a

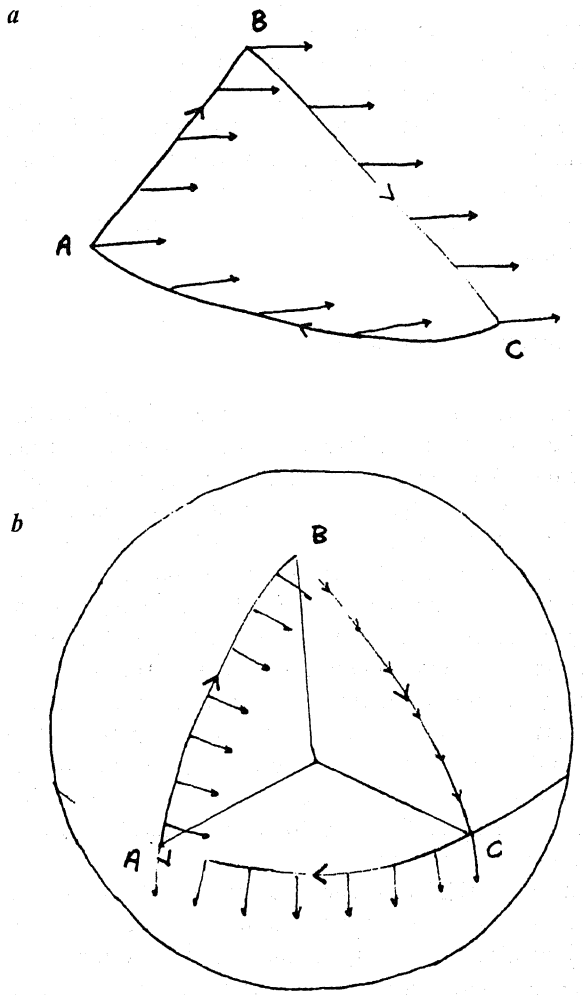


Figure 2. Extrinsic curvature: *a*, This is a parallel transport of a vector around a closed curve (ABC). The curve is drawn on a 2D sheet, embedded in 3D Euclidean space. There is no rotation of the vector, meaning there is no extrinsic curvature. *b*, Parallel transport of a vector along geodesics on the surface of a sphere, embedded in 3D Euclidean space. The vector is rotated 90°. The extrinsic curvature of a 3D spacelike hypersurface embedded in 4D spacetime is measured similarly. To obtain the curvature at a point on the surface, the closed curve (ABC) is pinched down infinitesimally.

three-dimensional Euclidean space. Since the cylinder can be obtained by rolling up a piece of paper, its intrinsic geometry is the same as that of the paper, i.e. a two-dimensional plane. The difference lies in how the paper and the cylinder sit in three-dimensional space. Technically, their extrinsic curvature is different.'

'Coming back to GR, the extrinsic curvature measures how the intrinsic geometry is changing in time. So, if we think of the intrinsic geometry as the configuration variable of general relativity, as one does in geometrodynamics, then extrinsic curvature can be thought of as the conjugate momentum. In quantum geometrodynamics, one would subject the two to the Heisenberg uncertainty relations.' (Figure 2)

By splitting spacetime into (3+1)D, are you treating the spatial and time coordinates on a different footing? Physically, this would correspond to choosing a special observer whose wristwatch keeps track of the distortions in the 3D spatial hypersurface. This method is very different to the 'covariant' schemes of quantization like the path integral approach.

'That is quite right. From aesthetic considerations, it would be nice to use, say, path integrals. However, no one has seen a way to make mathematical sense of path integrals in GR beyond certain simple minisuperspaces (which result when you freeze all but a finite number of degrees of freedom of the gravitational field and replace a quantum field theory problem by a quantum mechanics problem). The mathematical framework that my colleagues and I have developed over the last two years is, in principle, also applicable to the Euclidean path integral approach. However, there are two key difficulties. The first is conceptual. Even if one did have a complete Euclidean theory, it seems very difficult to extract physical predictions from it since, in quantum gravity, there is no analogue of the 'Wick rotation', which replaces the time coordinate t by it in Minkowskian quantum field theories, thereby enabling one to pass from the Euclidean to the Lorentzian regime (i.e. from Schwinger functions to the Wightman functions). The second difficulty is technical, associated with the specific form of Einstein's action.'

One of the main criticisms of the approach is that it breaks the general covariance of GR. Would the general covariance concept have any meaning if the continuum picture of spacetime is lost? One of the ideas brought up by your programme was the possibility of spacetime being combinatoric instead of geometric in nature, at the Planck scale.

'For technical clarity, I should first point out that, while canonical approaches do lack *manifest* covariance, they are, nonetheless covariant; for example, the full Poincaré group has a well-defined action on the phase space of electromagnetic fields in Minkowski space. With this side remark out of the way, let me answer your question. If anything resembling our approach is correct, then, in the quantum theory, there will be no such thing as a space–time geometry at a fundamental level. The picture, as you say, will be combinatorial and, therefore, the diffeomorphism group will not have a preferred role. Space–times will arise only semi-classically. In this approximation, one can ask if there is four-dimensional diffeomorphism invariance. The answer will be in the affirmative. However, it will be arrived at from the Hamiltonian picture; the four-dimensional covariance will not be manifested. This is a drawback of the approach but only from an aesthetic viewpoint.'

'Incidentally, for these reasons, I had first tried to use the covariant phase space for general relativity and

avoid the 3+1 splitting. That approach faced technical problems and I now believe that, even if the technical problems could be resolved, it would not lead to an interesting theory. I will not be surprised, however, if there is another way to ensure manifest covariance in the semiclassical regime.'

By casting GR into Hamiltonian form, it is found that the theory is a dynamically constrained system. These restrictions go by the name of the Hamiltonian and diffeomorphism constraints. What physical significance do they have?

'Constraints of this type signal the existence of a large gauge group. In electrodynamics, for example, there is an analogous constraint, the Gauss law, which says that the electric fields should be divergence-free. It ensures that the classical (and quantum) theory is gauge-invariant. Higgs pointed out quite early in the game that the diffeomorphism constraint of general relativity is rather similar. It ensures that the spatial diffeomorphisms – active coordinate transformations on the three-dimensional hypersurface – do not change physics. This constraint is universal in the sense that its form and meaning is the same for all gravity theories in which there is no background metric.'

'The Hamiltonian constraint is more tricky. It is responsible for time evolution, i.e. dynamics. If you treat it in the same fashion as the diffeomorphism constraint, one is led to what people call "the frozen (time) formalism", in which nothing evolves and physical states are represented by entire space–times. There is nothing wrong with this picture. It is like doing Hamiltonian dynamics in terms of constants of motion; it is, therefore, inconvenient from technical and interpretational perspectives.'

I have heard that the constraint equations are (high-order) polynomials in terms of the basic canonical geometrodynamical variables – the three-metrics and extrinsic curvatures. What effect does this have on trying to quantize the theory?

'It simplifies the technical problems of quantization enormously. In quantum field theories, what one might naively think of as operators are really operator valued distributions. Just as products of the Dirac delta distributions are ill-defined, products of these "operators" are also ill-defined. This is the origin of the infinities. So, if you have an expression which is non-polynomial in your basic operators, in general, there is very little chance that you would be able to regulate it and extract a meaningful operator. If the expression is a *low*-order polynomial, the chances of success are much greater. At various levels of rigour, this simplicity has been exploited by Rovelli, Smolin, Brügmann, Gambini, Pullin, Morales, Nicolai and Matschull to solve the quantum constraints. No such solutions exist in quantum geometrodynamics.'

'A simplification came in the mid-eighties when it was realized that there is another way of obtaining the Hamiltonian formulation where the basic variable is a connection which enables one to parallel transport chiral fermions. Thus, general relativity can also be regarded as connection dynamics. This formulation has two key advantages. The first is conceptual: general relativity is brought closer to gauge theories which govern all other basic forces of Nature. The second simplification is technical: equations of the theory become low order polynomials.'

The Ashtekar variables

At this stage, you came up with the idea of the Ashtekar variables. Where did the inspiration of these new variables come from? Was Amitabha Sen's work a major influence?

'For me, personally, a lot of the inspiration came from results on chiral solutions to the Einstein equations by Newman, Penrose and Plebanski in the mid-seventies. In a certain sense, this sector of Einstein's theory is completely integrable. Therefore, it seemed natural to try to base quantum theory on chiral variables. Sen's work was definitely an important influence up to a certain point. About the same time that Sen published his papers, Horowitz and I also had some results on positivity of the Hamiltonian of GR (for open universes) which used similar techniques. The two together suggested the direction that I finally took.'

What physical picture can you give us of them? You perform a canonical transformation of the old geometrodynamical variables, obtaining a connection and a triad as the new variables. A complex, group-valued, self-dual, spin connection is hard to visualize. In what way does a triad of orthonormal vectors on the three-slice take over the role of the extrinsic curvature?

'The connection enables us to parallel transport chiral fermions along closed loops and, therefore, has a direct physical meaning. Roughly speaking, the connection knows both about the (spatial derivatives) of the triad and the extrinsic curvature (the time derivative of the triad). If you think of the three-metric as the analogue of the position variable q in quantum mechanics and extrinsic curvature as the analogue of p, then the connection is analogous to the complex variable $z = q - ip$, which one often uses to construct the so-called coherent states in quantum mechanics. Clearly, if you know (q, p), you know (z, q) and vice versa. One can do quantum mechanics with (z, q) as basic variables, although it is somewhat unconventional. In quantum GR, we are led to connections and triads because these are the variables that simplify the equations.'

The actual canonical transformation is nonlinear. The result casts GR as a dynamical theory of connections. The phase space of GR is now seen to be embedded in that of Yang–Mills theory (a gauge theory). The constraints form surfaces which define restricted motions on

this phase space. This hidden relation seems remarkable. Were you expecting such a relation, or did it just drop out of the working?

'I knew that if the triads and connections can be shown to be canonically conjugate, the embedding you refer to would be a consequence. But it was far from obvious to me that the Poisson brackets between triads and connections would be so simple. Also, because of some subtle differences between the connections that Sen was using and the ones I was led to use, it was not obvious to me that all of Einstein's equations would be low order polynomials in the triads and connections.'

The next step is the transition from classical GR to quantum GR. By using these Ashtekar variables, you are trying to quantize a theory of connections, instead of a geometrodynamic theory based on spatial three-metrics. In terms of connections, the constraint equations of GR bear a great resemblance to those of *non-Abelian* Yang–Mills theory.

'Progress came about because there already existed a rich machinery to deal with quantum gauge theories. We did have to make some important changes eventually because in Yang–Mills theories, one makes a heavy use of the background Minkowski metric and our framework had to exist without reference to any background structures. But a number of key ideas came from Yang–Mills theories.'

'Incidentally, because all other basic interactions can be formulated in terms of connections, the geometrodynamic formulation did create a distance between GR and other gauge theories. For example, Weinberg, in the introduction to his book (see note 8) emphasizes this point. However, both Einstein and Schrödinger had presented a formulation of general relativity with connections, rather than metrics as the basic variables. The reason the idea did not catch on, I believe, is that their equations were even more complicated than those of geometrodynamics. They used affine connections and simplifications occur with chiral ones. I became aware of this piece of history only recently.'

Loop variables

In QM we have operators and states. Operators act on states, and take them from one allowed state to another. The simplest example of a QM system may be the simple harmonic oscillator. There the raising and lowering operators take a particle subject to a constant restoring force, to different energy eigenstates. The canonical variables, position and momentum, are raised to the status of operators, and they have to satisfy commutation relations. What are the corresponding operators and quantum states in this approach?

'This is a somewhat technical point and I hope you will bear with me. In particle mechanics, we have the notion of configuration space – the space of positions, for example – and quantum states are functions on this space. In field theories, the situation is more complicated. States are now functions on a *quantum* configuration space which is a genuine enlargement of the classical configuration space and the measures which dictate the inner product are typically concentrated on the exotic, nonclassical configurations.'

'The precise domain space of quantum states is a well-defined completion of the space of gauge equivalence classes of connections. This space can be constructed without reference to a background metric and we have also developed integral and differential calculus on this space to define Hilbert spaces and operators. There is no background metric or connection anywhere. This calculus tells us that there is a well-defined "nonlinear duality" between connections and loops and, hence, one can also represent states as suitable functions of loops. The connection and the loop representations are "dual" to one another in the same sense as the position and the momentum representations in quantum mechanics are dual. In practice, issues related to the Planck regime are often more transparent in the loop representation while semiclassical questions are generally easier to analyse in the connection representation.'

'For technical reasons connected with diffeomorphism invariance, quantum states are represented by objects which are more general than functions in the connection representation, while in the loop representation, one can regard them just as functions of loops. In either case the basic quantum operators are associated with loops and strips, i.e. ribbons. Their commutator algebra is surprisingly simple; commutators can be expressed in terms of rerouting of loops and strips, and intersections of loops with strips, and strips with strips. It is a pretty, geometrical picture.'

In the loop representation, quantum states arise as 'functionals' of Wilson loops. A deep relation exists between the equivalence classes of loops and the theory of knot classes, invariants and polynomials. There is also a relation to Chern–Simons theory.

'This comes about because there is a specific "nonlinear duality" between loops and connections: Given a closed loop and a connection, one obtains a number, the value of the trace of the holonomy of the connection around the loop. It turns out that this duality implies that there is a 1–1 correspondence between certain functions on the loop space and measures on the quantum configuration space. Therefore, we can take these loop functions as quantum states. One can then formulate the diffeomorphism constraint rigourously and seek its solutions. Not surprisingly, in the loop representation, the solutions are (certain types of) knot invariants. Thus, after you have imposed the diffeomorphism constraint, you can say that the quantum configuration space is the space of knot classes and all knot

invariants/polynomials are the potential quantum states. At a heuristic level, one expects the Chern–Simons theory to define a diffeomorphism invariant measure on the space of connections and these ideas have, in fact, been exploited by Gambini and Pullin to obtain solutions to the Hamiltonian constraint. Whether such results can be made rigorous is an open question; one would have to wait and see if the current work by mathematicians in this domain leads to useful results.'

Are you feeding in what the microstructure of space–time should look like beforehand, or does it come out of the equations? I have heard the analogy of Planck space–time being like a mesh of mail armour, made up of loops weaved together.

'A key point in our approach is that we do not start out assuming what quantum geometry should be like but let the theory lead us to it. (Of course, the framework as a whole makes certain assumptions – there should be no background structure and the Wilson loop operators should be well-defined – but they are of a general nature). We construct operators that correspond to geometry and ask for states which, when probed with these operators, would approximate a classical geometry at scales much larger than the Planck scale. We then obtain specific states which have this property but which display a discrete structure of a definite type at the Planck scale. These are the states that look like three-dimensional chain mail. The gravitational field is excited only at the loops in the chains. However, on coarse graining, the states are indistinguishable from classical geometries. More importantly, we find that spectra of basic operators such as the area of a given two-dimensional surface are quantized in the units of the Planck area.'

'In the loop representation, all operators act by breaking, rerouting and gluing of loops. These operations will code the quantum gravity interactions. So, in the Planck regime, the familiar Feynman diagrams will be replaced by "topological" diagrams representing these operations. Feynman diagrams will arise only in a low-energy regime, where, as I indicated above, the concept of gravitons is meaningful.'

What was Profs Smolin and Rovelli's inspiration to come up with the loop representation? Has there been any collaboration with the Trias and Gambini group, which is working on gauge theories?

'As I recall, they were trying to solve the diffeomorphism constraint and realized that the solution is "obvious" in the loop representation. At that time, we did not have the mathematical machinery involving measures on the quantum configuration space and it was hard to see how to construct solutions to this constraint in the connection representation. Therefore, their realization was a real breakthrough. Curiously, none of us knew about the work by Gambini and his collaborators at that time. One of the students in our group, Brügmann, came upon their papers while applying the "loopy ideas" to Yang–Mills theory and told us about them. We then contacted Gambini and Trias. Since then, Gambini, Pullin and others have made some very significant contributions. We are all extremely happy with these synergetic developments.'

You can also bring in supersymmetric matter couplings.

'Yes, supersymmetry can be brought in. There is work in this direction by Jacobson, Matschull, Nicolai and others. The work by Nicolai and his colleagues on three-dimensional supergravity has been especially illuminating; it showed that supergravity does allow an infinite number of physical states rather than just, say, a single solution analogous to the Hartle–Hawking wave function.'

Current work in the programme

Let us talk about some of the current research problems in the field. In the language of connections, you have to complexify the phase space of GR, since you have mixed (both complex and real) variables. Now you have to restrict the phase space to real portions, since the physical observables must be real quantities, not complex ones. In other words, you must implement reality conditions in the quantum theory. Have these been dealt with satisfactorily?

'Several special cases had been treated satisfactorily over the last few years. These include certain Bianchi models and midi superspaces with a single Killing field. Over the past year, we have been able to treat the *general case*. In the harmonic oscillator case, the Segal–Bargmann transform enables us to pass from the Schrödinger representation, where the wave functions $\psi(q)$ are complex valued functions of q, to the holomorphic representation, where they are *holomorphic* functions $\psi(z)$ of z. Since we know that q and p are self-adjoint in the Schrödinger representation and since the Segal–Bargmann transform is unitary, you know automatically that the reality conditions also hold in the holomorphic representation; you do not have to check it explicitly. We have constructed a rigorous analogue of this transform for quantum gravity. So, in principle, the problem has now been solved. However, we still have to develop calculational tricks to make the transform useful in practice.'

One of the main problems has been to define an inner product (vital to make sense of probabilities) in the Hilbert space. What is the status of this search?

'This problem arises at two levels: kinematical and dynamical. One generally needs an inner product on the "kinematical" states prior to imposition of constraints to make sure that the constraint operators are well-defined and that there are no anomalies in the constraint algebra. This part is taken care of by the recent mathematical developments. Furthermore, as I indicated above, the

diffeomorphism constraint has been regularized and solved. There are no anomalies. However, at the rigorous level, we still do not have a complete treatment of the Hamiltonian constraint. Very recently, the corresponding operator was constructed by Jerzy Lewandowski in Warsaw but we are yet to examine the question of anomalies. If there are none, we can construct a complete set of physical states. Fortunately, by now there are well-developed strategies to find the physically appropriate inner product and the question is if one of them will be easily implemented. However, it is quite possible that to ensure that there are no anomalies, we would have to bring in framed loops and quantum groups and/or supersymmetry. Mathematical machinery does exist to do this. And I would find it quite exciting if e.g. it is the dynamics of quantum gravity that ushers quantum groups into physics.'

'At a heuristic level, there has been a lot of progress on the issue of solutions to the Hamiltonian constraint and Rovelli and Smolin have introduced approximation techniques to find the inner product. These are useful guidelines for rigorous work. However, the problem of inner product on physical states is yet to be faced head-on.'

'Incidentally, if the programme is successful, we will have a quantum theory of gravity at a level of rigour that exceeds the current level in four-dimensional quantum field theories. This may seem like an overkill. However, since we have so little experience in nonperturbative QFTs, particularly in the absence of a background space–time, this level of care is needed both to ensure that we have not ignored some important subtleties and to convince the skeptics that a solution has really been found.'

How is the problem of time evolution handled? If there are only Hamiltonian constraints, and no Hamiltonian, how do you define a time variable, and evolve the quantum system?

'Two approaches have been followed, in both of which the notion of time is an approximate one. The first is some work I completed in 1989, where we saw that, in the connection representation, one can make a certain truncation of the theory in which the quantum Hamiltonian constraint can be written as the Schrödinger equation, where one of the connection components serves as time. This shows how the familiar evolution of QFT will arise in the low-energy regime. The second and more recent approach is due to Rovelli and Smolin. They couple gravity to a scalar field and use the scalar field as time. So, while we still do not know how to extract time and evolution in the Planck regime – and, like many of my colleagues, I feel this is not necessarily a meaningful thing to ask – in suitable approximations, time evolution has been recovered.'

For posterity, can I have your feelings as to where loop variables are heading?

'There are three parts to the answer. The first refers to quantum gravity, the second to other physical theories of connection, especially QCD, and the third to certain branches of mathematics.'

'Within gravity, as I indicated above, for the first time, we have a candidate for the quantum configuration space and the diffeomorphism constraint has been formulated without anomalies, and we know its general solution. If we can do the same for the Hamiltonian constraint – possibly by bringing in framed loops and quantum groups – one could say that quantum GR exists nonperturbatively. It is like saying, in QCD, that we have a representation in which the Hamiltonian is rigourously self-adjoint. This would be striking but not immediately solve problems of physical interest. The situation would be the same in QGR. However, that the theory exists, one could then look for reliable approximation methods and know that this is a meaningful programme. Many such methods are already being developed. I believe that it is crucial to continue to look for such methods.'

'Since the mathematical machinery that has been developed works for any theory of connections, it is tempting to apply it to QCD. We have already done so in two space–time dimensions and obtained several new results. The most notable among these are closed expressions for the Wilson loop Schwinger functions and a proof of equivalence of the Hamiltonian and Euclidean path integral methods. One might venture further and try to extend these methods to three and four space–time dimensions. That they will lead to mathematically interesting quantum field theories is clear. Whether these theories will be as useful in physics is an open issue. In any case, it is exciting that one and the same methods are being applied to theories of all basic forces of Nature.'

'On the mathematical side, there have also been some interesting new results. A few years ago, it was widely believed that spaces of gauge equivalent connections would not admit *any* nontrivial diffeomorphism-invariant measures. Several of us have found infinite families of such measures by now. The relation between knots and these diffeomorphism-invariant measures is equally exciting. Our methods also enable one to extend (sufficiently nice) invariants of ordinary knots to those of generalized knots, where loops are allowed to intersect and can have kinks and overlaps. This is of interest to knot and graph theorists. Such results arise because we are looking at the familiar structure from a new angle, with a new perspective. Such contributions, I believe, will continue.'

Superstrings[22–25]

Perhaps the most popularized of all the recent approaches to QG is superstrings.

Superstring theory proposes that the elementary constituents of matter are one-dimensional curves, rather

than point particles. According to the theory, the quarks, leptons and gauge bosons of the standard model (SM) arise as excitation states of a truly fundamental entity – the superstring. These excitations come from rotational or vibrational degrees of freedom (rather like the harmonic modes in a violin string), or from internal degrees of freedom like supersymmetry (SUSY) and Lie group symmetries. If they exist, superstrings would be the smallest things in nature. They would be about 10^{-33} cm long, and reside in the weird world of Planck scale physics. Superstring excitations energies are about 10^{18} GeV, or 10^{-5} g. These particle masses are way too big to be detected by current accelerators, including the large hadron collider (LHC) being built in Europe. We are looking at detecting particles as massive as bacteria!

Superstrings can come in two varieties. One resembles a line segment with free ends. This type is called the open superstring. The other type is a loop with the topology of a circle – the closed superstring. Both versions reside in a ten-dimensional (9+1) spacetime.

Since we only see four (3+1) dimensions in real life, it is assumed that the six extra spatial dimensions are curled up, in a region as small as the Planck scale. The process of curling up is called compactification. There is no mechanism in the theory to tell how you curl the extra dimensions up, and this arbitrariness is a cause for concern. Theorists have explored Calabi–Yau spaces, group manifolds, and other spaces as candidates. The idea of compactification is reminiscent of Kaluza–Klein theory, which attempts to unify gravity with electromagnetism using an extended spacetime of 5 or (4+1) dimensions. This theory was extensively expanded in the 1970s to include the Yang–Mills theory, which gives for example, the type of non-Abelian gauge theory used in weak interactions (this involves a spacetime of dimension greater than 5; e.g. 12 for SU(3) internal symmetry). The mathematical techniques of these larger-dimensional spacetimes (including what is known as 'dimensional reduction' – the means whereby one gets back to the physical dimension 4) have been much used in discussions of supergravity and superstring theory. However, it is important to emphasize that these developments all work at the level of the *classical* field equations: the quantum theory used in quantising them is essentially standard.

String theories have an extraordinarily rich mathematical structure. Exploring their symmetries, described by various algebras, has been a major focus of work in the field. It has emerged that the same mathematical structures appear in many different fields. For example, as it moves through spacetime, the string traces out a world sheet or cylinder. Like the world lines (geodesics) of point particles, this world surface is constrained to be of extremal area. The symmetry properties of these surfaces are similar to those encountered in two-dimensional condensed-matter systems. There has been valuable cross-fertilization between these fields.

Superstring theory is also related to subjects like 2D conformal field theories, nonlinear sigma models, WZW models and other integrable models. It has associations with 2D supergravity, Lie superalgebras, and W-algebras, the latter being world-sheet symmetry algebras of bosonic string theories. Quantization leads to problems with 'anomalies' in these algebras, whereby classical symmetries are violated at the quantum level. Work is still being done on anomalies in W-algebras and also more generally in Lie algebras.

The vibrational spectrum of open superstrings includes a massless spin-1 gauge boson associated with a Yang–Mills group – SO(32) or $E_8 \times E_8$. These groups are needed for anomaly cancellations. For the closed superstring, a massless spin-2 particle arises. It has been identified as the graviton, the conjectured carrier of the quantum gravitational force.

GR is taken to be the large-distance or low-energy limit to the superstring. Strings seem like point particles at weak energy scales (10^{-16} cm) since it is many orders of magnitude less than the superstring scale (10^{-33} cm). A small extended object looks like a point from the perspective of lower energies. (Figure 3)

By examining string interactions, it is found that the existence of open strings implies closed strings. Closed strings alone can form a consistent system. It follows that every consistent string theory necessarily includes gravity, since closed strings have gravitons in their excitation spectrum.

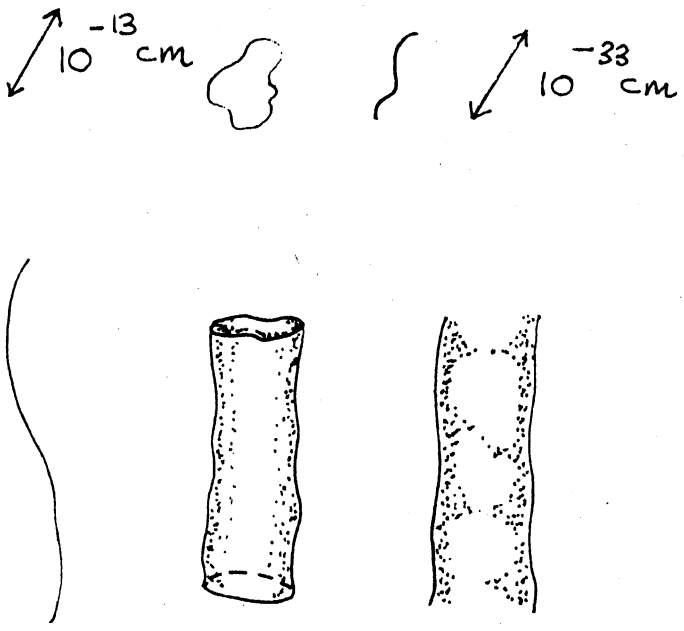
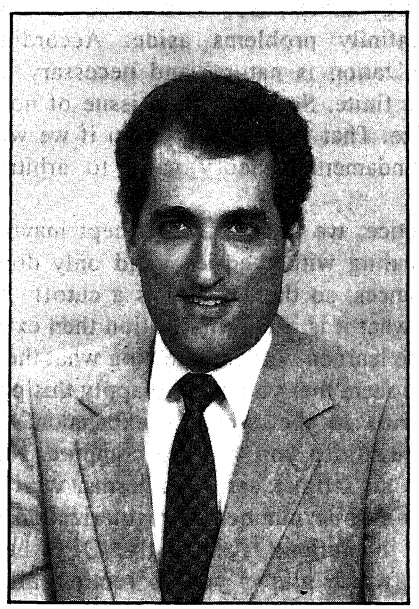


Figure 3. From particles to strings. Point particle theories describe physics well at weak scale energies, but it is possible that at Planck scales, particles are actually tiny strings. World lines would then be replaced by world cylinders and sheets. See *Superstrings*, Vol. 1, M. Green, J. Schwarz and E. Witten, Cambridge University Press, 1987.

People have tried to generalize strings to higher-dimensional objects, like P-branes. However, it is very hard to make quantum theories of extended objects and get them to satisfy physical consistency constraints like causality and unitarity. In fact, the more extended you go, the harder it is. The success of string theory relies on conformal symmetry, which is a property solely of 2D world sheets.

The self-consistency of superstrings is heavily dependent on two experimentally undetected phenomena: supersymmetry and higher spacetime dimensions.

Ed Witten is at the Princeton Institute of Advanced Studies, New Jersey. He is a major contributor to superstring theory. He was once a history major at Brandeis, but had some science background. His interest in superstrings was sparked by a review article by John Schwarz. He says: 'Schwarz's review article made it easier to start learning about developments in superstring theory than it had been before – but my concentrating on reading it was, of course, a result of the fact that I was already very curious about superstrings.'



Edward Witten

Quantum field theory began when Planck tried to explain the emission spectrum of black body radiation. He postulated that energy was emitted discontinuously in discrete units, or quanta. Later, from a statistical analysis of the radiation field, Einstein suggested it was not just the emission and absorption of radiation which was quantized, but also the actual electromagnetic field itself – in the form of photons. This idea was confirmed experimentally by the Compton effect. Naturally, there were attempts to extend the quantum field idea to the classical gravity field, described by a metric tensor in Einstein's GR. The basic excitation of the quantum gravity field would be spin-2 radiation, which results from the symmetric nature of the metric tensor. How hard would it be to detect the gravitational analogue of the photon – the graviton – given the weakness of the gravitational interaction? Ed replies: 'Classical gravitational waves (for which there is indirect evidence from timing measurements of the binary pulsar) could be directly detected by laser interferometers – perhaps eventually giving also an important new window into some exotic phenomena in astrophysics. It is also possible that a very long wavelength cosmic background of gravitational waves could be detected by timing measurements involving space ships and pulsars. These experiments could indeed possibly verify directly the spin-2 nature of gravitational waves.'

'Unfortunately, there is no experiment in sight that could directly verify the quantized nature of the gravitational field.'

The gravity field is different from the other force fields since it plays a dual role. On the one hand, it defines the arena in which other fields interact. Yet, gravity itself is a participant in the interaction. Does QFT assume that spacetime is flat to all energy scales, i.e. has a Minkowskian spacetime signature even at Planck scales? 'QFT as we know it, without quantizing gravity, works on a fixed space–time manifold, which for most obvious purposes we can take to be Minkowski space. There is then a continuum space–time down to arbitrarily small distances.'

Gravity is highly nonlinear since the gravity field couples to mass–energy. The gravitational field gravitates. Is this a major conceptual subtlety when we try to quantize gravity? 'There are some conceptual issues in quantizing gravity, but it is not at all clear that they are decisive. What to me is the decisive difficulty is the problem of the infinities that arise when one tries to quantize the usual theory.' (It is instructive to contrast this point of view with Abhay's, stated earlier on.)

Renormalization, infinities and anomalies

QFT works with perturbation expansions around a Minkowskian background. Terms in these expansions are represented diagrammatically by Feynman diagrams. The Feynman diagram has an attractive interpretation as a representation of an actual physical process, and (though limited to perturbation theory) this picture has some validity.

In QED interactions between photons and electrically charged particles are linear. However, the situation is different with gravitons. Since energy gravitates, two gravitons can bend each others path. Gravitons can also emit and absorb other gravitons and couple to the other gauge bosons.

QFT introduces the concept of a particle's self-energy. Let us take the electron self-energy as an example. According to QFT, there is a difference between a

bare electron and a physical electron. A bare electron is an unphysical entity because it has no radiation field around it. Such a radiation field is known to exist because of Heisenberg's uncertainty principle, which allows for the presence of a blanket of virtual photons. When a bare electron interacts with this virtual field, the bare electron is converted to a physical electron. This self-interaction affects the energy of the bare electron, and hence its mass. Quantities like charge are affected as well. In fact, all physical quantities are affected by the self-interaction.

When the Feynman integrals are evaluated for these self-energy terms, the integrals diverge to infinity. A similar phenomenon occurs when we calculate the photon self-energy in the form of vacuum polarization diagrams.

The first order terms (tree level) are finite, but the higher-order terms (the radiative corrections) involving closed loops diverge. Physically, the difficulties arise as a result of the point particle concept. Heisenberg's principle says that energy conservation can be violated locally such that the relation $\Delta E \cdot \Delta t > \hbar$ is obeyed. Since an electron is taken to be a point, emission and reabsorption of a virtual photon need take no time at all. So the virtual photon can have an infinite amount of energy. This contributes an infinite mass to the physical electron, since it must haul around the virtual photon blanket. Other properties of the electron get affected similarly.

Despite these shortcomings, QFT is one of the most successful theories we have. On this Ed says: 'Just because we write down a quantum field theory Lagrangian containing some fields does not mean that we know what the particles are going to be. QCD is an example where the particles (mesons and baryons) do not even appear in the Lagrangian (which contains quarks and gluons). Even if the particles are in natural one-to-one correspondence with the fields, as in QED, the masses of particles cannot be just read off from the constants in the Lagrangian. It is necessary to compute all kinds of quantum corrections. The fact that the nature and masses of the particles are, in large part, computed and not postulated, is a manifestation of the fact that the QFT description of nature is more thorough than its predecessors.'

To handle these divergences, QFT invokes the principle that energy cannot be measured, only differences in energy can be. A scheme called renormalization was developed by Feynman, Tomonaga and Schwinger back in 1947 to take care of this, for QED.

To handle divergent graphs, we firstly impose cut-offs on the integrals. This modifies the theory into another one. The cut-offs set some energy scale limit to the theory's predictions. Ed comments: 'The cutoff theory has physically unacceptable properties, so it is necessary to remove the cut-off and "renormalize". This is just as well, since the cut-off is arbitrary and if one did not have to remove it one would lose most of the predictive power. In good cases, like QED, after renormalization the parameters in the quantum theory are in one-to-one correspondence with the classical parameters; thus, in QED the only parameters are the charge and mass of the electron (the mass can be scaled to one by a convenient choice of units).'

The second step is to renormalize the theory by incorporating the blanket of virtual particles around the bare particle. This converts the bare particle into a physical one. In the process, relations between bare and physical quantities are introduced. These relations contain renormalization constants. Thirdly, we revert from the renormalized theory to the original QFT by taking the limit when the cut-offs go to infinity. This restores the original theory. The divergences reappear, but only in the renormalization constants relating physical and bare quantities. Like the bare quantities, the constants are not observable. The original integrals, written in terms of physical quantities, become well-defined and finite to all energy scales! So, in a sense we have shifted all the infinity problems aside. According to Ed: 'Renormalization is natural and necessary even if the effects are finite. So there is no issue of not having to renormalize. That would be so even if we were dealing with a fundamental theory valid to arbitrarily short distances.'

'In practice, we are always (except maybe in string theory) dealing with theories valid only down to very short distances, so there really is a cutoff – we just do not know what it is. Renormalization then extracts whatever can be learned without knowing what the cutoff is.'

It would seem that you have to apply this procedure to all the terms in the S-matrix expansion, for all the physical quantities you want to compute. Now the expansion is an infinite series of terms, which seems to require an infinite number of regularization and renormalization operations. Amazingly, QED has a gauge invariance which allows you to remove all the divergences from the measurable quantities to all orders of perturbation expansion, via just one type of renormalization.

Why cannot we compute physical quantities directly, instead of going via this roundabout? Feynman himself thought that renormalization was 'sweeping the problem under the carpet'. He thought that we should be able to formulate a theory such that we get zero ground state energy from the very start. What physics could we be missing here? One possibility, already mentioned by Abhay, is that the infinities ultimately arise from neglecting effects in Planck scale physics which would have a subtle effect on low energy physics.

When we try to fit gravity into QFT, the algorithm is to linearize the GR field by choosing a coordinate system in which the metric tensor can be expressed as

$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, where $\eta_{\mu\nu}$ is the Minkowski metric and $h_{\mu\nu}$ is a small perturbation such that $|h_{\mu\nu}| << 1$. This small perturbation, a tensor field, is subject to quantization on a Minkowski background which provides the causal structure. It is treated as the 'graviton' particle.

'Formally, there is no problem in constructing the perturbation expansion of the S = matrix along the above lines. The only real difficulty is that one meets ultraviolet divergences that cannot be renormalized away (without introducing new parameters absent in the classical theory).'

At the one-loop level pure QG's divergences can be renormalized away. However, the divergences arising at the two-loop level cannot be removed without robbing the theory of predictive yield.

On the associated problems of anomalies in QG he comments: 'The existence of a quantum theory is always delicate. Once one writes down a classical Lagrangian, one does not automatically get a quantum theory. Infinities and anomalies are two of the main ways to fail. Anomalies that affect gauge invariance are even worse than lack of renormalizability; they mean that the perturbation expansion of the quantum theory just does not make sense.' Essentially, the time rate of change of some quantity is zero in the classical theory and upon quantization this rate becomes nonzero, breaking some symmetry and, hence a conservation law.

Recent work in superstrings

How does extending the idea of particles to superstrings change things? An analogy would be helpful here. Let us take the decay of the free neutron, which is mediated by the weak force. We know that the old Fermi theory of weak decay is not renormalizable. However, the electroweak theory is. The electroweak theory replaces the point interaction of the old Fermi theory by inserting a massive W boson propagator, which mediates the decay. The theory is now renormalizable. Virtual supermassive superstring excitation states affect the renormalizability of superstring interactions by 'softening' or spreading out the interactions.

Ed says: 'Certainly, in going to string theory, everything becomes softened, and eventually finite. How best to understand it is not yet clear, but somehow the ultraviolet region is missing in string theory.'

Although the superstrings are extended objects, their interactions still take place at single spacetime points. However, this point depends on the Lorentz frame of the observer (the pants diagram phenomena) unlike the point particle case. There the spacetime interaction point is identical in all Lorentz observer frames. Thus there is less freedom in constructing string theories compared to the point particle case. This is why we have only a finite number of superstring theories instead of the infinite number of point field theories. (We have to choose the correct point particle QFT via phenomenology.) (Figure 4)

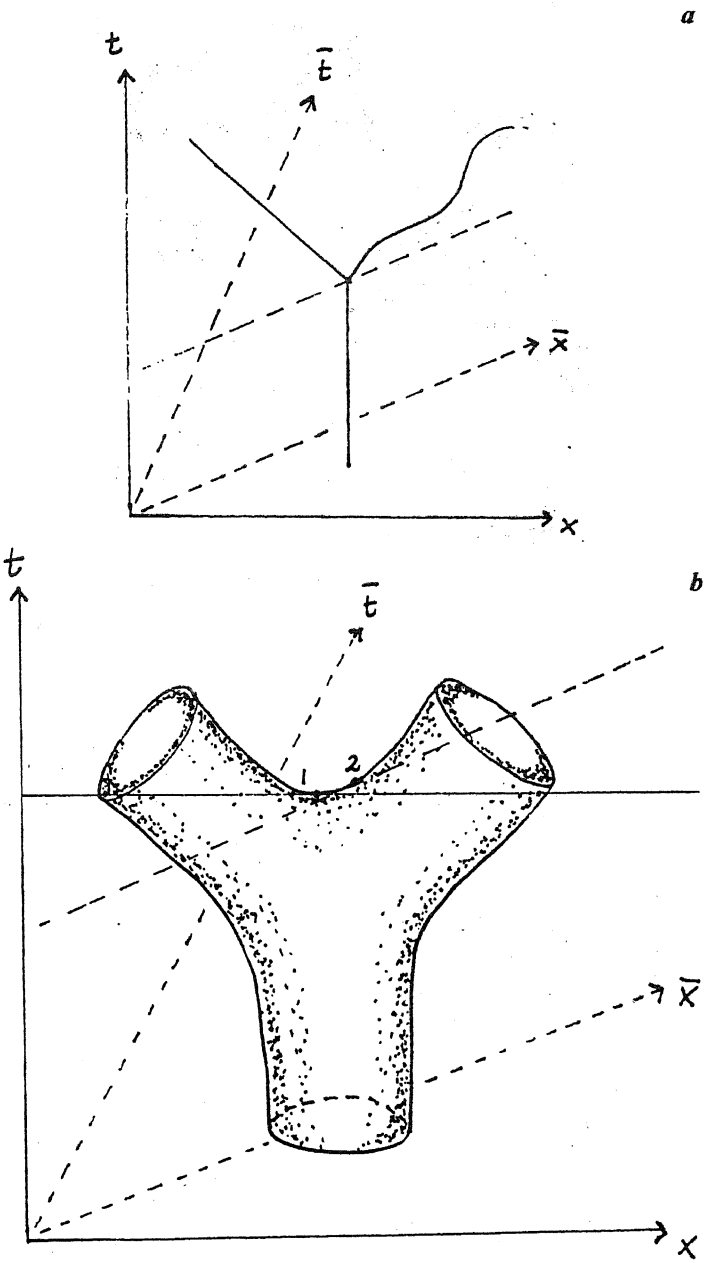


Figure 4. Contact interactions for particles and strings. *a*, Interactions between particles occur at a single spacetime point, called the vertex in the Feyman diagram approach. This point is Lorentz-invariant since it is the same in all reference frames. *b*,. The splitting of a closed string into two closed strings, otherwise known as the pants diagram. The world surface is a cylinder, so the interaction region is no longer a point. One observer will see the interaction occur at point (1), and another will see it occur at (2). The contact point is no longer Lorentz-invariant. Divergences in QFTs of point particles can arise when two sharp vertices get arbitrarily close to one another. Renormalization is needed to remove the ultraviolet divergences by introducing ultraviolet momentum cutoffs, thereby 'fuzzying' out the vertices. Since string diagrams are surfaces, they tend to accomplish this process automatically. See *Quantum Field Theory of Point Particles and Strings*, B. Hatfield, Addison-Wesley, 1992.

According to one counting scheme, there are currently three d=10 superstring theories (type I, type II, heterotic) and although they have no adjustable dimensionless parameters, there are many mathematically consistent solutions which seem to satisfy the superstring equations of motion. Choosing physically representative solutions phenomenologically seems no different to inserting SM parameters by hand. The problem of finding physical solutions can only be solved if we know how compactification works. Doing this we can find the ground (vacuum) state of the theory. From there, we can work out the low-energy excitation spectrum, find the ratios of the excitation masses, and determine couplings constants from the topology of the compactified space.

Ed is optimistic about this arbitrariness: 'The attractive near-uniqueness that string theory appears to possess in principle is largely lost in practice because one finds a plethora of possible classical solutions. Hopefully, in time we will learn something new – perhaps connected with the question of why the cosmological constant vanishes – and the nature of the issue will change.'

On the problem of formulating a nonperturbative theory of superstrings, Ed says: 'My work with Nathan Seiberg (and related work he did by himself) led to some new results about nonperturbative behaviour of some quantum field theories, but not yet to new results on nonperturbative behaviour of string theory.'

String theory appears to admit a group of discrete field transformations, called S-dualities, as exact nonperturbative quantum symmetries. How could these symmetries help in formulating a nonperturbative QG theory of strings?

'I don't know yet, but I think it is a very exciting clue. S-dualities are symmetries that have no real analogue in our previous experience in physics (they do persist in some field theories – including some of those I studied with Seiberg – which are related to low energy limits of string theories).'

Particles trace out world lines, but strings trace out world sheets or cylinders, which can be treated as Riemann surfaces. When the Feynman diagrams are worked out for the closed string, it is found that they become finite order by order in the genus expansion.

Ed says: 'In string theory everything is finite – there is no need for renormalization.'

'Feynman diagrams are replaced by Riemann surfaces; these have much additional symmetry and structure, and this is related to some of the known wonders of the subject – there are surely also unknown wonders (S-duality is a hint of them), but since they are unknown I cannot tell you about them.'

On the matter of progress on the physical principles behind superstrings which should generalize Einstein's equivalence principle, he replies: 'The new physical principles haven't yet been found. It is hard to say how much progress there has been, since some of the things that have been learned may well turn out to be important clues.'

Superstrings are supposed to reside in ten spacetime dimensions, but we are really only speaking approximately. The phase space of the superstring consists of all the possible orientations and states the string can be in. What are the problems of formulating superstrings directly in four spacetime dimensions from the start? Are not superstrings supposed to generate spacetime, instead of just residing in it, since superstrings have gravitons in their excitation spectrum?

'It is true that in our present understanding ten spacetime dimensions is the classical solution of string theory with the maximum symmetry. Exactly how strings can be so smart as to generate space–time is still a mystery, but they do.'

'You can formulate strings "directly in four dimensions from the start" but then you do not have the degree of uniqueness that one has in ten dimensions. That is why the theory is often said to be ten-dimensional.'

There is no mechanism in superstring theory which tells us how the extra dimensions should compactify. Feynman said about superstrings: 'I don't like it that they're not calculating anything. I don't like it that they don't check their ideas. I don't like it that for anything that disagrees with experiment, they cook up an explanation – a fix-up to say "Well, it still might be true". For example, the [superstring] theory requires ten dimensions. Well, may be there's a way of wrapping up six of the dimensions. Yes, that's possible mathematically but why not seven? When they write their equation, the *equation* should decide how many of these things get wrapped up, *not* the desire to agree with experiment ....' What is your response to this remark?

'Some of the predictions are indeed checked. For instance, strings predict gravity while prestring physics makes quantum gravity impossible. This prediction has been tested – and the result is one of the main reasons for the interest in string theory.'

Are you familiar with the Ashtekar canonical gravity programme? If you are, do you think that the programme is founded on valid physical assumptions?

'Some of the work in that programme is interesting to me. I'd personally be surprised, however, if it turns out that conventional general relativity exists as a nonperturbative quantum field theory.'

Do you think the LHC will shed any light on low-energy superstring phenomenology? The detection of supersymmetry partners could be crucial. To explain the absence of SUSY partners at low energies we have introduced the idea of the spontaneous breaking of supersymmetry, analogous to the Higgs mechanism for electroweak symmetry breaking: one facet of the SM which is unconfirmed is this mysterious Higgs field which pervades all spacetime, which makes the electroweak

theory work by giving masses to particles, and which we have no shred of evidence for.

'Apart from gravity, supersymmetry is one of the general predictions of string theory. Therefore, confirming it at the LHC would give a big boost to the whole effort. Moreover, if supersymmetry is observed, it won't be just that experimentalists will say "Yes, there is supersymmetry". They will observe a plethora of superpartner masses and couplings. There is a reasonable hope that in all that data there will be clues to how string theory should be done – or tests of theoretical ideas that may have emerged in the meantime.'

I asked Ed what the inspiration for his ideas was. Did he manipulate page-long equations in his head? Did he deal with concepts, and then phrase everything in maths later on? Where Einstein, Galileo and Newton were different from everyone else was they could ask really simple questions which had shattering consequences for existing physical theories. Was this Ed's modus operandi too?

To this he replied: 'I am afraid I don't work on the cosmic scale you are suggesting. Most of the time I am bogged down doing nothing, and whatever success I have has come by focusing on a very little bit of the total picture.'

Outlook

It is clear from the discussion that we are a long way off from even the foundations of a full theory of QG. Despite some famous claims, the end of theoretical physics is still not in sight. I asked Chris about possible lines of future investigation.

'I think that the superstring programme and the Ashtekar programme are currently the only two really viable approaches to trying to construct a full quantum theory of gravity, and I, therefore, support strongly work on both fronts!'

'Myself, I have long been bothered by the use of continuum ideas in both quantum theory and classical general relativity, whereas both programmes do take this for granted. So I would like to be able to get away from this, but the problem is to find a theoretical framework that does not seem to be contrived and artificial. One of the many reasons why the superstring and Ashtekar programmes *are* so important is that they do not have this contrived nature: within their own terms, each has quite compelling reasons for taking it seriously.'

'Similarly, I (like Roger Penrose) believe in my heart that quantum theory itself needs to be changed radically to reflect our changing conception of space and time. But, again, it is difficult to do this in a way that is believable to one other person than the author of the paper (Salam's definition of what it means to say that a theory is worth taking seriously!).'

Ed Witten gives this advice to those who want a try at superstrings: 'It is hard to give general advice to young researchers. But people who want to learn strings should definitely make sure that along the way they get a thorough knowledge of what is known in particle physics and how it can be (more or less) derived from strings.'

As to where superstrings are heading he has this to say: 'Lots of really nice things have been discovered, and that will continue. Sooner or later, probably far in the next century, the underlying unifying ideas will be discovered and then the physical landscape will be very different from what we can imagine today. If we are lucky and work hard, maybe we will live to see the day!'

Further reading

1. Isham, C. J., in *Canonical Relativity: Classical and Quantum*, (eds Ehlers, J. and Friedrich, H.), Springer, Berlin, 1994, pp. 1–21.
2. Isham, C. J., in *Recent Developments of Quantum Fields* (eds Mitter, H. and Gaustener, H.), Springer, Berlin, 1992, pp. 123–225.
3. Isham, C. J., in *Canonical Relativity, Classical and Quantum* (eds Ehlers, J. and Friedrich, H.), Springer, Berlin, 1994, e-Print Archive: gr-qc/9310031.
4. Isham, C. J., in *Integrable Systems, Quantum Groups, and Quantum Field Theories* (eds Ibort, L. A. and Rodriguez, M. A.), Kluwer Academic Publishers, London, 1993, pp. 157–288.
5. Isham, C. J. and Linden, N., Continuous histories and the history group in generalized quantum theory, Imperial-TP-94-95-19, Mar 1995, 24 pp, e-Print Archive: gr-qc/9503063.
6. Brügmann, B., Bibliography of publications related to classical and quantum gravity in terms of the Ashtekar variables, SU-GP-BIB, March 1993, 14 pp, pre-Print Archive: gr-qc@xxx.lanl.gov-9303015.
7. Brügmann, B., Loop representations, in *Bad Honnef 1993, Proceedings, Canonical gravity*, pp. 213–253 and München MPI Phys, MPI-Ph-93-094 (93/12, rec. Dec.), 38 pp, pre-Print Archive: gr-qc/9312001a.
8. Kuchar, K. V., Canonical quantum gravity, UU-REL-92-12-10, n.d. (recd April 1993) 40 pp, e-Print Archive: gr-qc@xxx.lanl.gov - 9304012.
9. Ashtekar, A., Recent mathematical developments in quantum general relativity, CGPG-94-9-4, n.d. (recd Oct 1994) 13 pp, e-Print Archive: gr-qc@xxx.lanl.gov - 9411055.
10. Ashtekar, A., in *Canonical Gravity from Classical to Quantum* (eds Ehlers, J. and Friedrich, H.), Springer, Berlin, 1994, pp. 327–367.
11. Ashtekar, A., Quantum Gravity: A Mathematical Physics Perspective, in The Proceedings of the International Conference on Mathematical Physics Towards the 21st Century, Negev, Beer Sheva, Israel, 14–19 March 1993; eds Sen, R. N. and Gersten, Beer-Sheva University Press, 1994.
12. Ashtekar, A., Mathematical Problems of Nonperturbative Quantum General Relativity, SU-GP-92-11-2, Dec 1992, 87 pp. Invited talk at Lectures given at Les Houches Summer School on Gravitation and Quantizations, Session 57, Les Houches, France, 5 July–1 August 1992. e-Print Archive: gr-qc@xxx.lanl.gov -9302024.
13. Ashtekar, A. and Stachel, J., eds, Conceptual Problems of Quantum Gravity. Proceedings, Osgood Hill Conference, North Andover, USA, 15–18 May, 1988, Birkhauser, 1991, pp 602. *Einstein studies*, vol 2.

14. Ashtekar, A., *Lectures on Nonperturbative Canonical Gravity*, World Scientific, Singapore, 1991, pp. 334 (Advanced Series in Astrophysics and Cosmology, vol 6).
15. Ashtekar, A., New Variables for Classical and Quantum Gravity, *Phys. Rev. Lett.*, 1986, **57**, 2244–2247.
16. Smolin, L., Experimental signatures of quantum gravity, CGPG-95-3-2, Mar 1995, e-Print Archive: gr-qc/9503027.
17. Rovelli C. and Smolin, L., Spin networks and quantum gravity, CGPG-95-4-4, April 1995, 42 pp, e-Print Archive: gr-qc/9505006.
18. Rovelli, C. and Smolin, L., Discreteness of area and volume in quantum gravity, CGPG-94-11-1, Nov 1994,. 36 pp, e-Print Archive: gr-qc/9411005.
19. Ashtekar, A., Lewandowski, J., Donald, M., Jose, M. and Thiemann, T., Quantization of diffeomorphism-invariant theories of connections with local degrees of freedom, UCSBTH-95-7, Apr 1995, 71 pp, e-Print Archive: gr-qc/ 9504018.
20. Ashtekar, A. and Lewandowski, J., Differential geometry on the space of connections via graphs and projective limits, CGPG-94-12-04, Dec 1994, 53 pp. e-Print Archive: hep-th/9412073.
21. Ashtekar, A., Lewandowski, J., Donald, M., Jose, M. and Thiemann, T., Coherent state transforms for spaces of connections, CGPG-94-12-02, Dec 1994, 38 pp, e-Print Archive: gr-qc/9412014.
22. Cumrun, V. and Edward, W., A one loop test of string duality, HUTP-95-A015, May 1995, 13 pp., e-Print Archive: hep-th/9505053.
23. Schwarz, J. H., Classical duality symmetries in two dimensions, CALT-68-1994, May 1995, 13 pp., Presented at STRINGS 95: Future Perspectives in String Theory, Los Angeles, CA, 13–18 March 1995, e-Print Archive: hep-th/9505170.
24. Schwarz, J. H., String theory symmetries CALT-68-1984, March 1995, 12 pp., presented at International Workshop on Planck Scale Physics, Puri, India, 12–21 December 1994, e-Print Archive: hep-th/9503127.
25. Schwarz, J. H., Evidence for nonperturbative string symmetries, CALT-68-1965, July 1994, 9 pp., Conference on Topology, Strings and Integrable Models (Satellite Colloquium to the ICMP-11 18–23 Jul 1994), Paris, France, 25-28 July 1994, e-Print Archive: hep-th/9411178.

Notes

1. Ontology is the part of metaphysics concerned with the nature of existence.
2. Stephen Hawking's *A Brief History of Time: A Reader's Companion*, prepared by Gene Stone, Bantam Books, 1992, p125.
3. In SR, a foliation is a one-parameter family of three-dimensional *flat* surfaces (hyperplanes) which, taken all together, fill out the entire spacetime. In GR, these surfaces are curved.
4. Dust as a standard of space and time in canonical quantum gravity. By J. David Brown (North Carolina State University), Karel V. Kuchar (Utah U.), GRQC-9409001, Aug 1994. 57 pp.e-Print Archive: gr-qc@xxx.lanl.gov - 9409001.
5. Dust is a basic type of matter which experiences only gravity, and no other interaction.
6. *Physics Today*, 1991, **44**, 36, *Decoherence*, Wojicek Zurek.
7. See for instance, *Extremal Black Holes and Elementary String States*, Ashoke Sen (ICTP, Trieste and Tata Inst.). TIFR-TH-95-19, Apr 1995. 16 pp. e-Print Archive: hep-th/9504147.
8. *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity*, Steven Weinberg, New York, Wiley, 1972, xxviii, 657 pp., ill, 23 cm.
9. A single-valued function $f(z)$ is holomorphic (or regular or analytic) in a region if it is differentiable at each point of the region.
10. Superstring Theory, *Phys. Rept.*, 1982, **89**, 223.

ACKNOWLEDGEMENTS. This article was based on interviews with Chris Isham, Ed Witten and Abhay Ashtekar. William Spence at Queen Mary and Westfield College in London helped review the final draft. It was the basis for a theory seminar presented at Melbourne University, Australia.

Faint star counts and the Milky Way structure

Devendra Ojha

The Milky Way Galaxy offers a unique opportunity for testing theories of Galaxy formation and evolution. Here we discuss how large surveys, both photometric and astrometric, of galactic stars are the keystones in investigations into such fundamental problems as the merging history and future of the Galaxy. This work features a sample survey plan to produce probes of stellar populations in the Milky Way. Objectives of this work are to trace the fine structure of our Galaxy through the statistical study of the stellar distributions according to their luminosity, colours and proper motions. This involves two steps: first, acquiring a new photometric and astrometric sample survey in various galactic directions; and second analysing the data using a model of population synthesis and determining the properties of populations in the Galaxy and constraints on the scenario of formation and evolution.

THE field of galactic structure is currently extremely active, due mostly to improvements in observational capabilities and the realization that our Galaxy is perhaps the one best-suited to test theories of Galaxy formation. The concept of stellar populations has proved to be one of the most useful ideas of modern astronomy. Most simply, it is the idea that we can define a population as a set of stars that possess shared characteristics such as composition, age or kinematics, and that we can use the properties of various stellar populations to de-

Devendra Ojha is in the Inter-University Centre for Astronomy and Astrophysics, Post Bag 4, Ganeshkhind, Pune 411 007, India.

termine the structure and evolutionary history of the Galaxy. The first modern definition of stellar populations was given by Baade[1] who defined two stellar populations – the now famous populations I and II – based on colour-magnitude diagrams and on the brightness of resolved stars in M31 and M32. Population I is composed of stars found in the disk of our Galaxy and in the Magellanic Clouds; a signature of this population is the presence of highly luminous O- and B-type stars and open clusters. Population II is made up of stars with colour–magnitude diagrams like galactic globular clusters, which lack luminous blue stars; another signature of population II is the presence of short-period Cepheid variables. Now it became clear that population II was old and poor in metals, while population I was metal-rich and contains both young and old stars. The various doubts and worries about the population I/II scheme came to a head in the famous Vatican Conference[2] of 1957, during which Baade's original two populations were replaced with five: extreme population I, older population I, disk population, intermediate population II, and halo population II. Since the Vatican Conference, the field of stellar populations has been advanced significantly, and we now have a wealth of detailed information on populations throughout the Galaxy. One of the main developments is that we now have a good picture of the intermediate population (thick disk) in our Galaxy, i.e. a component with characteristics intermediate to those of the thin disk and halo populations[3].

Studies of stellar populations often take as their starting point large-scale surveys, with selection criteria involving one or more observable variables. The physical location of the stars is arguably the most important constraint. With these points in mind, we have carried out a sample survey in UBV photometry and proper motions in various directions in the Galaxy. Large Schmidt telescope plates are well suited for measuring photometric magnitudes and positions for a large number of stars, because they have a large field ($5° \times 5°$) and go sufficiently faint. By choosing a reasonable number of fields in various galactic directions, we expected to get a representative sample of the spatial distribution of the different stellar populations, which have led to an understanding of the history of formation and evolution of the whole Galaxy. The kinematical studies of stellar populations are very important because they follow the galactic dynamics and retain the memory of the evolution of the whole potential. The results discussed here concern the kinematics of thin disk and thick disk populations of our Galaxy and the scenario of formation for the thick disk population.

Observations and data reduction

The chosen directions constitute a complete set of fields at high and intermediate latitudes:

- Direction towards the galactic anticentre[4] ($l = 167°$, $b = 47°$).
- Direction towards galactic centre[5] ($l = 3°$, $b=47°$) near the globular cluster M5.
- Direction towards the galactic antirotation[6] ($l = 278°$, $b = 47°$).
- Direction towards the North Galactic Pole[7] ($l = 58°$, $b=80°$)

The basic observational material for the survey is Schmidt plates. The plates used in our programme have been taken from Palomar, Tautenburg, European Southern Observatory and Observatoire de la Côted'Azur (France) Schmidt telescopes. For each field the plates have been taken in UBV filters. In all we have a total of 58 plates for the various fields. Each plate has been scanned with the MAMA (Machine Automatique à Mesurer pour l'Astronomie) machine at Paris Observatory. A sampling step and pixel size of 10 μm has been used to scan the plates. The output of the MAMA machine is a catalogue of objects giving the positions X and Y, integrated density and the area of each object. The integrated densities are used to calibrate the plates photometrically, while a combination of various Schmidt plates of a given field is used for proper motion determination involving a time base of ~34 years. To calibrate the plates photometrically a number of photometric standards are required. The standards should cover the entire range of magnitudes and colours to be studied. To obtain photometric standards we have observed a number of subfields in each field using the CCD system attached to the 1-m Sampuranand telescope of the UP State Observatory at Nainital, India. A few standards have also been observed using the 1.2-m telescope of Observatoire de Haute-Provence in France. Our final photometric and astrometric survey contains 55,000 stars covering a 50 square degree field.

Kinematical analysis

The distance of each star in the catalogue was determined by estimating the absolute magnitude, which was obtained from a M_v vs $B–V$ relation of main-sequence stars of solar metallicity, taking into account the metallicity change as a function of the distance from the galactic plane. The cardinal components of the stellar space velocity (in km/s), U, V and W (where U is defined as positive in the direction of the galactic anticentre, V is positive in the direction of galactic rotation, and W is positive in the direction of NGP), were derived from proper motions, μ_l and μ_b (in arcsec per year) and line of sight distance d (in pc)[8].

The major factors contributing to distance errors are the photometric errors in the colour of σ_{B-V} and magnitude σ_V. From our surveys, we expect the mean errors in

$\sigma_{B-V} = 0.1$ and $\sigma_V = 0.07$ for $V = 11$–18. From these values, our estimate of error of about 20% in distance seems to be realistic up to $z = 3$ kpc, where z is the distance above the galactic plane.

Kinematics of stellar populations

The kinematical properties of each stellar population are related to their spatial distributions. Scale height, velocity dispersions and asymmetric drift are linked by the Boltzmann equation. The proportion of each stellar population varies with the distance above the galactic plane and the selection of a stellar sample at a given distance allows one to optimize the proportion of one population. Since kinematical data allow one to improve this identification of populations, we have minimized bias from mutual contamination of each population at a given height by performing a kinematical separation of populations. To perform the kinematical separation, we have used a maximum-likelihood method (Stochastic Estimation Maximization (SEM) algorithm)[9] in order to deconvolve the multivariate Gaussian distributions and estimate the corresponding parameters. The aim of the SEM algorithm is to resolve the finite mixture density estimation problem under the maximum-likelihood approach using a probabilistic teacher step. Through SEM one can obtain the number of components of the Gaussian mixture (without any assumption on this number), its mean values, dispersions and the percentage of each component with respect to the whole sample. We have used this method to separate the 2-D Gaussian velocity distributions (U, V) to identify the two components (thin disk and thick disk) of the Galaxy[10]. A multivariate discriminant analysis[11] is also used to distinguish the thick disk from other populations with the help of the model of population synthesis[12–14].

Thin disk population

For the thin disk population, we observe a continuous increase of velocity dispersion with height[4,5]. It can be explained by the fact that the thin disk is a mixture of disks with various scale heights and velocity dispersions, since the disk is not formed in isothermal components. This results in a vertical gradient of velocity dispersions, since the disk is not resolved in isothermal components. For the thin disk population, we obtain a measure of the kinematic gradients, which appear to be

$$\frac{\partial \ln \sigma^2_{U+W}}{\partial R} = -0.18 \pm 0.03 \text{ kpc}^{-1}$$

and

$$\frac{\partial \ln \sigma^2_{V}}{\partial R} = -0.05 \pm 0.02 \text{ kpc}^{-1}.$$

R is the galactocentric distance projected upon the galactic plane. Our results[10,11] confirm that the thin disk has a relatively short scale length of ~2.3 ± 0.5 kpc and scale height of ~258 ± 50 pc.

Thick disk population and the merging history of the Milky Way

The thick disk population has been revisited under the light of these new data. We used also other photometric and astrometric sample survey in complementing direction, like the pole[7].

The sample of stars in two fields (galactic centre and anticentre) have been divided into 6 or 7 bins of distance, and in each bin of distance a fit has been performed with a SEM algorithm to identify the Gaussian (U, V) components corresponding to the different populations[4,5]. The thick disk population has been identified as a discrete and distinct component. In our present study[11], the thick disk population appears to be isothermal, since the kinematical characteristics are nearly constant along the line of sight and so that there is no vertical gradient. For this reason, the velocity characteristics for this population are much better defined. For the thick disk population, the kinematic radial gradients appear to be almost zero:

$$\frac{\partial \ln \sigma^2_{U+W}}{\partial R} = -0.008 \pm 0.01 \text{ kpc}^{-1}$$

and

$$\frac{\partial \ln \sigma^2_{V}}{\partial R} = +0.02 \pm 0.02 \text{ kpc}^{-1}.$$

By combining the kinematical results deduced from our four surveys (galactic centre, anticentre, antirotation, and NGP), we have derived the velocity ellipsoid of the thick disk population[11]. The mean kinematic parameters are: (σ_U, σ_V, σ_W, V_{lag}) = (72±4, 54±3, 40±3, 40±10) km/s. The most probable values of scale height and local normalization for the thick disk component are determined to be h_Z ~759±50 pc and $A_{\text{thick}} = 7.4^{+2.5}_{-1.5}$% of the disk. The ratio of the number of thick disk stars in the galactic centre region to that in the centre region yields h_R ~ 2.8–3.7 kpc for the scale length of thick disk[11].

These properties have been compared with several scenarios of formation. 'Top-down' models, where the thick disk forms through a dissipational collapse, after the halo formation and before the thin disk has completely settled down, are in contradiction with the data, since they infer a continuity in the characteristics of the thick disk and thin disk and large kinematical vertical gradients. Alternatively, 'bottom-up' models suppose a formation of a thick disk after the complete collapse of the thin disk. We can eliminate those where the thick

disk formed from a secular kinematic diffusion of thin disk stars (which would produce a continuity between thin and thick disks). It remains a scenario well in agreement with the present data: the thick disk could have been formed from the dynamical heating of the thin disk during the sink of a small galaxy into the Milky Way. This merging event has to happen at the beginning of the thin disk life time so that the gas can cool again and form stars in the long-lasting thin galactic disk that we see now. This violent bottom-up scenario leaves two important observational signatures. First the thick disk is a separate population distinct from the thin disk and the halo. Second, no gradient can be generated in the thick disk by the event, although a preexisting gradient may survive the merger.

Conclusions

The thick disk population is found from all points of view (density laws, kinematics and metallicities) as a population well separated from the thin disk. It shows no gradient either on abundance[3] or on kinematics. The results emerging from the present study of the correlations between photometry and kinematics give a mounting evidence that the thick disk of the Galaxy is most likely a sequel of a dwarf satellite galaxy merging in the Milky Way disk during the early epoch of this disk. This interpretation emerged from an accurate characterization of the thick disk properties: the scale height and the scale length of this component have been established; its rotation and velocity dispersion turned out to be quite distinct from both the disk itself and the halo. The local density of the thick disk component is twice what was previously assumed and the stellar colours do not reflect any significant chemical gradient.

1. Baade, W., *Astrophys. J.*, 1944, **100**, 137–146 .
2. Blaauw, A., in *Galactic Structure* (eds Blaauw, A. and Schmidt, M.), Univ. of Chicago Press, Chicago, 1965, pp. 435–453.
3. Robin, A. C., Haywood, M., Crézé, M., Ojha, D. K. and Bienaymé, O., *Astron. Astrophys.*, 1995 (in press).
4. Ojha, D. K., Bienaymé, O., Robin, A. C. and Mohan, V., *Astron. Astrophys.*, 1994, **284**, 810–826.
5. Ojha, D .K., Bienaymé, O., Robin, A. C. and Mohan, V., *Astron. Astrophys.*, 1994, **290**, 771–784.
6. Ojha, D. K., Bienayme, O., Mohan, V., Haywood, M. and Robin, A. C., *Astron. Astrophys.*, 1995 (in preparation).
7. Soubiran, C., *Astron. Astrophys.*, 1992, **259**, 394–403.
8. Murray, C. A., *Vectorial Astronomy*, Adam Hilger, Bristol, 1983, pp. 271–311.
9. Celeux, G. and Diebolt, J., *Rev. Stati. Appl.*, 1986, **34**, 35–52.
10. Ojha, D. K., Ph D thesis, Strasbourg University, France, 1994, pp. 75-83.
11. Ojha, D. K., Bienaymé, O., Robin, A. C., Crézé, M. and Mohan, V., *Astron Astrophys.*, 1995 (submitted).
12. Robin, A. C. and Crézé, M., *Astron. Astrophys.*, 1986, **157**, 71-90.
13. Bienaymé, O., Robin, A. C. and Crézé, M., *Astron. Astrophys.*, 1987, **180**, 94-110.
14. Haywood, M., Ph D thesis, Paris Observatory, France, 1994, pp. 28-69.

ACKNOWLEDGEMENTS. I thank Drs O. Bienaymé, A. C. Robin, M. Crézé, and V. Mohan for useful suggestions and helpful discussions during the course of this work. Financial assistance from IFCPAR, New Delhi, is also thankfully acknowledged.

REVIEW ARTICLE

A unique and remarkable binary pulsar

Simon Johnston

Research Centre for Theoretical Astrophysics, University of Sydney, NSW 2006, Australia

PSR B1259–63 is one of the most remarkable pulsars ever discovered. It is the only pulsar known to be orbiting a main-sequence Be star and forms the 'missing link' in the evolutionary scenario of binary stars. Its orbit is highly elliptical and every 3.5 yr it approaches to within 0.5 AU of its companion star. Its pulse profile is unique among pulsars and shows that the opening angle of its emission cone exceeds 180°. Observations made around the time of closest approach of the two stars show changes in the pulsar's dispersion measure, rotation measure and fractional linear polarization on very short timescales. The pulsar is observed to spin down during the periastron passage. The companion star to the pulsar has been monitored at optical and UV wavelengths. The system is a radio and X-ray transient and has been detected at γ-ray energies, with the high-energy emission probably occurring in a shock front formed by the collision of the two stellar winds.

IN this review paper, I will discuss the remarkable binary pulsar PSR B1259–63. I will start with a general introduction into pulsars and binary systems before moving on to the specifics of the system containing PSR B1259–63. There are an increasing number of excellent conference proceedings and review articles on pulsars in general with varying degrees of complexity. Interested readers should consult refs. 1–11 and the references therein.

Almost 30 years after the discovery of the first pulsar by Hewish *et al.*[12], radio pulsars remain one of the astronomy's most exciting fields. The existence of neutron stars was predicted as early as the 1930s by Baade and Zwicky[13], who speculated that when a high-mass star (typically 5 or more solar masses, $M_\odot$) ends its life in the spectacular event known as a supernova explosion, it should leave behind a dense stellar core. Although a blue star at the centre of the Crab Nebula was postulated to be such a neutron star[14], it was little thought that they would be detectable. In 1967, however, these neutron stars were serendipitously discovered by a group of radioastronomers working at Cambridge University in the UK. The name pulsars (a contraction of the two words *pulsating stars*) is a misnomer; the stars rotate rather than pulse and the principal characteristic of a pulsar is its incredibly stable rotation rate. The radioemission originates from near the magnetic pole and the misalignment between the rotation and magnetic axes produces a regular pulse of radiation for an observer on Earth (like the oft-quoted analogy with a lighthouse).

Pulsars are now known to be highly magnetized, rapidly rotating neutron stars which originate from the supernova explosions of high-mass stars. Pulsars have masses in a tight range around 1.4 $M_\odot$ and have diameters of only ~20 km; this implies a stellar density in the region of 10^{17} kg m^{-3}. Almost 600 have now been discovered and catalogued[15]; all bar a few are located within our own Galaxy. Their spin periods range between 0.00155 and 5.4 s and their inferred magnetic fields lie between 10^4 and 10^9 T. The youngest known pulsar is 900 years old in the Crab Nebula. The oldest pulsars have ages in excess of 10^9 years, comparable to the age of the Galaxy. All pulsars are observed to spin down with time due to rotational energy losses from relativistic particles streaming from the high magnetic field.

Only ~2% of pulsars are part of a binary system. However, their importance is not proportional to their number and they provide many valuable insights into the physics and the formation and evolution of neutron star systems and their progenitors. Before the discovery of PSR B1259–63, all the companions to known binary pulsars were degenerate stars, either white dwarfs or other neutron stars. The evolutionary route of double neutron star systems is thought to occur as follows[7]: Two high-mass stars are in a mutual orbit; the more massive of the two stars evolves first and then explodes in a supernova explosion. In rare cases, the system remains bound and consists of a high-mass star and a young neutron star[16]. The high-mass stars eventually evolve, and the system goes through a high-mass X-ray binary phase (HMXB) when material from the evolved star accretes onto the neutron star surface. Eventually, the high-mass star explodes in its turn and, if the system remains bound, a double neutron star system remains. All the steps of this evolutionary chain had been observed (most famously perhaps for the double neutron star system PSR B1933+16, for which the discoverers[17] won the 1993 Nobel Prize for Physics) except the second step – the neutron star, high-mass star pair before it evolves into an HMXB. This 'missing link' was found with the discovery of PSR B1259–63 (ref. 18) and an important piece in the evolutionary jigsaw fell into place. A further example of such a system has subsequently been discovered[19,20].

It turned out that PSR B1259–63 was orbiting a $10M_\odot$ star known as a Be star (described in more detail below). Be stars[21] themselves are rather strange objects. Most stars show spectral lines only in absorption; however, the Be stars are characterized by having the Balmer series of hydrogen in emission. In the early 1930s, this was interpreted to mean that the star had a disk of material surrounding its equator which had been thrown off due to centrifugal forces from the rapidly rotating underlying star[22]. This simple model is still prevalent today, although observations, particularly at UV wavelengths, have found that this cannot be the whole story. There remains no convincing theoretical model of Be stars. The general picture, however, is that Be stars have a hot, tenuous, fast-moving wind which is thought to originate from the stellar poles. They also have a disk of material which is cool, dense and slow-moving and which follows Keplerian orbits at least for a few tens of stellar radii. In the simplest model the disk can be characterized by two parameters: an opening angle and a radial power law distribution of electron densities within this envelope[23].

Discovery

PSR B1259–63 was discovered during a search for young, distant pulsars in the galactic plane using the Parkes 64 m telescope located in NSW, Australia[24]. For a variety of reasons, the search was conducted using an observing frequency centred at 1520 MHz rather than the 400 MHz commonly employed in pulsar searches. This high-frequency turned out to be rather valuable as, most unusually for pulsars, PSR B1259–63 cannot be detected at low frequencies in short integration times.

The pulsar was interesting to us right from the moment of discovery. It had the shortest spin period (47 ms) of any of the pulsars discovered in the survey and had a rather curious double-pulse profile reminiscent of the archetypal young pulsar in the Crab Nebula (PSR B0531+21). After a successful attempt to measure the period change over a few days, we concluded that the pulsar was probably young. This encouraged us to look for a supernova remnant around the pulsar, but to our disappointment an image of the region with the Molonglo Synthesis Telescope (MOST) revealed

nothing of interest in the pulsar's vicinity.

By the middle of 1990 we realized that the pulsar was not behaving exactly as it should. In two consecutive timing observations the pulsar was extremely weak before reappearing about a month later at its usual strength. As the timing programme continued and more data points were obtained, it was evident that the pulsar's behaviour could not be explained with a simple period and period-derivative, and in early 1991 we came to the conclusion that the pulsar was a member of a long-period binary system. The most remarkable thing about the binary companion is that it had to have a mass greater than 2 $M_\odot$ and this implied that the companion was either a black hole or a high-mass star. Both possibilities were extremely interesting – all other binary pulsars to that date, of which there are anyway very few, had either white dwarf companions (with mass less than 1 $M_\odot$) or neutron star companions (with masses between 1.2 and 1.6 $M_\odot$).

A sub-arcsecond position using the Australia Telescope Compact Array (ATCA) was obtained and it turned out that a well-known 10th magnitude Be star, SS 2883, was located within the ATCA error box. Thus, the companion was a main-sequence star with a mass of ~10 $M_\odot$, the first such system discovered and an important system for our understanding of the evolutionary sequence of the millisecond pulsars. When the paper announcing the discovery of the system was published[18], not enough data had been obtained to allow an accurate determination of the orbital parameters; indeed, it turned out that the errors quoted in the parameters were much too optimistic. However, the general shape of the orbit was correct; it was highly eccentric ($e \sim 0.9$) and the binary period was in excess of a few years. This remains the orbit with the highest eccentricity out of all the known neutron star binary systems. This high eccentricity implies that the pulsar comes very close to its companion star at closest approach ('periastron').

Waiting for periastron

Although by 1992 it was apparent that periastron was still a year or more away, there were plenty of experiments needing to be done. The programme to monitor the pulsar in the radio was continuing, with data being obtained roughly every two weeks. As the time baseline increased, the binary parameters became better determined, and by early 1993 we were confident of the numbers and published them in an IAU telegram[25] with the prediction that the next periastron would occur on 9 January 1994.

At optical wavelengths, a crude spectrum of the star had been obtained[18], but in April 1993 we used the 3.8 m optical Anglo-Australia Telescope (AAT) to obtain high-resolution spectrum of the prominent hydrogen Balmer lines typical of Be stars as a class. The spectrum showed the higher-order Balmer lines and the He I line

Table 1. Observed and derived parameters for PSR B1259–63 (from ref. 27)

Right ascension, α (J2000)	$13^h\ 02^m\ 47^s.68(2)$
Declination, δ (J2000)	−63° 50′ 08″.6(1)
Period, P	47.762053919(4) ms
Period derivative, $\dot{P}$	$2.2793(4) \times 10^{-15}$
Epoch of period	MJD 48053.44
Dispersion measure, DM	146.75(8) cm^{-3} pc
Orbital period, P_b	1236.79(1) days
Projected semimajor axis, $a \sin i$	1295.98(1) light s
Longitude of periastron, w	138.6548(2)°
Eccentricity, e	0.869836(2)
Epoch of periastron	MJD 48124.3581(2)
Magnetic field, B	3.3×10^{11} G
Characteristc age, τ_c	0.33 Myr
Mass function, $f(M_p)$	1.53 $M_\odot$
Predicted periastron passage	MJD 49361.2 ≡ 1994 Jan 9.2.

in emission and thus confirmed that the star was around spectral type B1 and had a mass of ~10 $M_\odot$ and a radius of ~6 $R_\odot$. This, in turn, implies that the inclination angle of the binary orbit in the plane of the sky is roughly 35°. The optically derived distance of only 1.5 kpc is in contrast to the distance derived from the dispersion measure of 4.5 kpc (ref. 26); it is likely that the optical value is closer to the truth. The equatorial stellar rotation velocity is ~180 km s^{-1}; given the star's spectral type, the velocity is expected to be somewhat higher than this. If we ascribe the velocity difference to an inclination angle effect then the equator of the star is also inclined at 35° to the line of sight, similar to the binary orbit. Although

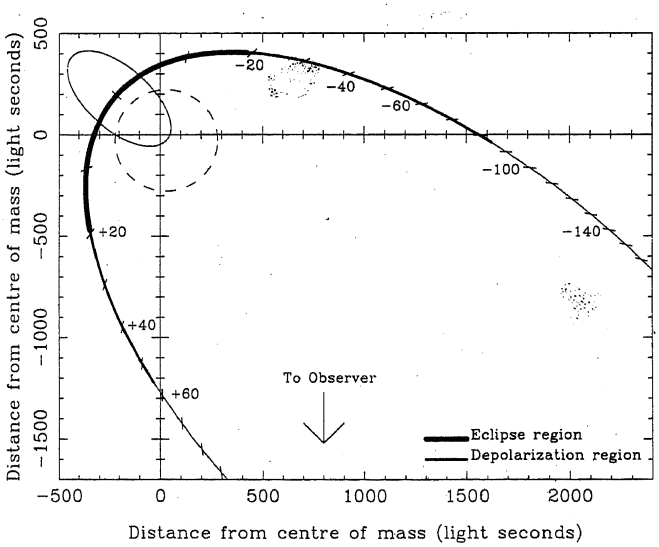


Figure 1. Schematic diagram showing the orbital path of the pulsar and its Be star companion around the time of periastron. Tick marks on the pulsar orbit are at 10-day intervals. The orbit of the Be star, assuming a mass of 10 $M_\odot$ for this star, is shown in the top-left corner, and the likely size and position of the optical circumstellar disc at periastron is shown by the dashed circle. The medium-thicknes line indicates the portion of the orbit over which the pulsed emission was depolarized, and the full-thickness line indicates the interval over which the pulsar was eclipsed. (After ref. 40)

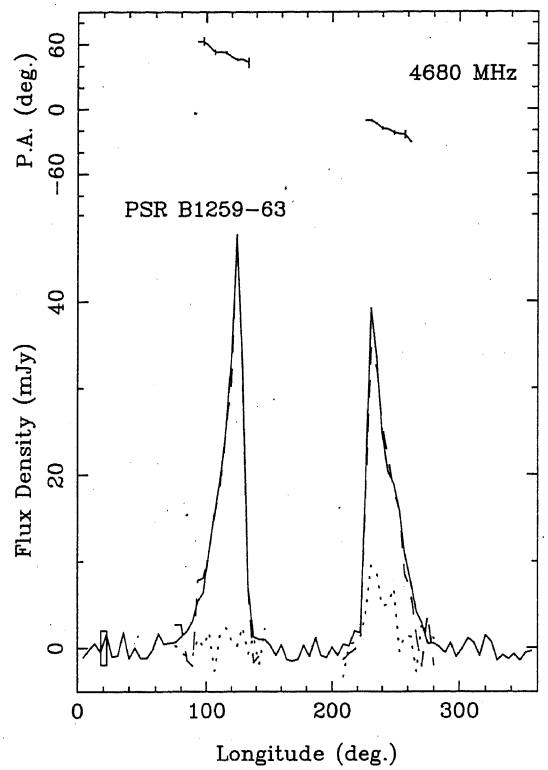


Figure 2. Mean pulse profiles and polarization parameters for PSR B1259–63 at 4.7 GHz. In the lower part, the solid line is the total intensity (Stokes parameter I) profile, the dashed line is the linearly polarized intensity (L), and the dotted line is the circularly polarized intensity (V). In the upper part, the position angle of the linearly polarized part (ψ) is plotted where L is significantly above noise, with $\pm 2\sigma$ error bars on every second point. An error box of amplitude 2σ and width equal to the effective profile resolution (including dispersion smearing across frequency channels) is visible at the left-hand end of the profile baseline. (After ref. 28)

this result is highly model-dependent it does give some indication that the stellar disk and the binary orbit are close to aligned. A combination of the optical results and the improved timing parameters for the pulsar using all the data available up to mid-1993 was published by Johnston *et al.*[27]. The observed spin and Keplerian parameters and a few of the derived parameters of the system are shown in Table 1. A schematic of the pulsar's orbit near the time of periastron is shown in Figure 1.

Radio polarization data for the pulsar were obtained at frequencies ranging between 1500 and 8400 MHz. The data show that both pulses are strongly (80%) linearly polarized at all frequencies[28]. The polarization properties of the pulsar at 4700 MHz are shown in Figure 2, which shows the pulsar in its 'conventional' profile, and this leads one to believe that the pulse shown on the left arrives at the observer before the pulse shown on the right. Remarkably, the polarization data show that convention in this case is flaunted. The data show that the pulsar emission cone must have a very large opening angle (greater than 180°) and that the pulse shown on the right actually arrives first[28]. The data also confirm that the binary inclination angle must lie close to 35°.

Observations had also begun at higher energies. Following an apparent nondetection of the pulsar just prior to apastron at X-ray wavelengths[29], the ROSAT telescope was used to observe the pulsar at two epochs just after apastron[30]. The pulsar was detected on both occasions, the second detection being roughly double the flux of the first. There was no evidence for any pulsed flux at either epoch. Cominsky *et al.*[30] considered various origins of the X-rays and ruled out that they originated from the Be star's hot coronal wind or from Roche-lobe overflow accretion. They concluded that some accretion from the Be star wind must be occurring even at such a large distance from the star. Later, the apparent nondetection was reanalysed[31] and it was found that the pulsar was, in fact, detected at a low flux level; however, this did not alter the basic conclusions in ref. 30.

Meanwhile, the theorists had not been idle. Kochanek[32] was the first out of the starting blocks and he modelled the impact of the pulsar on the equatorial disk of the Be star and the observational consequences for the optical emission lines. King and Cominsky[33] had taken the X-ray observations and proposed that the X-rays originated from low-level accretion from the tenuous (at apastron) Be star wind spherically onto the neutron star. In the model, the equatorial wind speed had to be low, only ~10 km s^{-1} and relatively cool. Later X-ray observations (see below) probably rule out this explanation. Lipunov *et al.*[34] proposed an evolutionary scenario for the pulsar and also predicted that changes in the pulsar's dispersion and rotation measures would be observed at periastron. Finally, shock theory was used [35] to model the interaction between the pulsar wind and the Be star wind and implications of the theory for high-energy observations near periastron were discussed. The conclusions were that the observations should yield a determination of the physical characteristics of both winds.

Periastron

Observations at the time of periastron were carried out from the radio through to γ-ray energies, encompassing UV, optical and X-ray energies.

Careful timing analysis of the pulsar for over 1500 days and two periastron passages came to fruition with the discovery that the pulsar had spun down slightly as a result of its periastron passage[36]. The magnitude of the spin-down is rather small, with a $\Delta P/P$ of $\sim 2 \times 10^{-9}$ and it is argued that this is a result of the so-called 'propeller

effect'[37,38]. The slowdown occurs because although the strong magnetic field of the pulsar prevents accretion onto its surface, the incoming material still interacts with the strong magnetic field at the Alfvén radius. The magnetic field 'propels' the material away from the star but at the same time the material imposes a braking torque on the pulsar. Detailed modelling of the effect in this system has been considered by Ghosh[39].

Observations of the pulsar using the Parkes radiotelescope over a few months around periastron have been discussed by Johnston *et al.*[40]. The observations showed changes in the pulsar's dispersion measure, rotation measure, flux density and percentage polarization on short timescales. This is the first time such changes have been observed in any pulsar. Essentially, the pulsar acts as an excellent probe of the wind surrounding the Be star. As the pulsar approaches its companion, it enters denser regions of the Be star's wind and the pulsar's dispersion measure starts to increase. Eventually, due to the clumpy nature of the wind, the pulses themselves become scatter-broadened and this renders the pulsar undetectable at low frequencies. Finally, a combination of pulse scattering and free–free absorption renders the pulsar undetectable for a period of 6 weeks around periastron itself[40]. On egress from the eclipse, the rotation measure and percentage polarization of the pulses varied rapidly over a period as short as a few hours. Melatos *et al.*[41]. looked at detailed modelling of the stellar environment based on the radio observations of the pulsar. They concluded that the pulsar wind cannot dominate the Be star's wind contrary to popular belief and that the best fit to the observed radio data comes from assuming that the Be star has a thick equatorial disk with a exponential density profile. The value of the stellar magnetic field and surface density are consistent with other values found in the literature[42].

Observations of the binary system were made with the ASCA X-ray satellite on three epochs around periastron[43]. The system was detected on all three occasions with an X-ray luminosity of $\sim 10^{34}$ erg s^{-1} just before and just after periastron and about a factor of 2 lower at periastron itself. As with the earlier ROSAT data, no pulsations were detected and accretion onto the neutron star surface as a cause of the X-rays can also be ruled out[43]. The most likely explanation is that magnetohydrodynamic shock acceleration processes between the pulsar and the Be star's wind are the most likely explanation for the X-ray emission. Details of this mechanism will appear in a later publication[44].

Finally, observations at γ-ray wavelengths with GRO[45], in the UV[46], optical[47] and the radio continuum[48] are all still being analysed and the results should appear in the press within this year.

Summary

I have given a brief overview of the discovery of the peculiar binary pulsar PSR B1259–63. It remains one of only two pulsars known to be in orbit around a high-mass main-sequence star. It is the only pulsar to show dynamical changes in its measured parameters and the only radio pulsar to show X-ray and γ-ray transient phenomena.

Most of the observational data are currently in press or in preparation (as of July 1995) and the publication of this material will undoubtedly stimulate further research into this system. The next periastron will occur in mid-1997 and once again both observers and theoreticians will be watching with interest.

1. Taylor, J. H. and Stinebring, D. R., *Annu. Rev. Astr. Ap.*, 1986, **24**, 285–327.
2. Helfand, D. J. and Huang, J., *The Origin and Evolution of Neutron Stars, IAU Symposium No. 125*, Reidel, Dordrecht, 1987.
3. Srinivasan, G., *Astron. Astrophys. Rev.*, 1989, **1**, 209–260.
4. Ögelman, H. and van den Heuvel, E. P. J., *Timing Neutron Stars*, Kluwer, Dordrecht, 1989.
5. Lyne, A. G. and Smith, F. G., *Pulsar Astronomy*, Cambridge University Press, 1990.
6. Backer, D. C. and Kulkarni, S. R., *Physics Today*, 1990, Mar, 26–35.
7. Bhattachary, D. and van den Heuvel, E. P. J., *Phys. Rep.*, 1991, **203**, 1–124.
8. Ventura, J. and Pines, D., *Neutron Stars: Theory and Observation*, Kluwer, Dordrecht, 1991.
9. Hankins, T. H., Rankin, J. M. and Gil, J. A., *The Magnetospheric Structure and Emission Mechanisms of Radio Pulsars, IAU Colloquium 128*, Pedagogical University Press Zielona Góra, Poland, 1992.
10. Bailes, M. and Johnston, S., in *Review of Radio Science 1990–1992* (ed. Stone, W. R.), Oxford University Press, Oxford 1993, pp. 677–714.
11. Phinney, E. S. and Kulkarni, S. R., *Annu. Rev. Astr. Ap.*, 1994, **32**, 591–639.
12. Hewish, A., Bell, S. J., Pilkington, J. D. H., Scott, P. F. and Collins, R. A., *Nature*, 1968, **217**, 709–713.
13. Baade, W. and Zwicky, F., *Proc. Natl. Acad. Sci.*, 1934, **20**, 254–259.
14. Baade, W. and Zwicky, F., *Phys. Rev.*, 1934, **45**, 138.
15. Taylor, J. H., Manchester, R. N. and Lyne, A. G., *Astrophys. J. Supp. Series*, 1993, **88**, 529–568.
16. Bailes, M., *Astrophys. J.*, 1989, **342**, 917–927.
17. Hulse, R. A. and Taylor, J. H., *Astrophys. J.*, 1975, **195**, L51–L53.
18. Johnston, S., Manchester, R. N., Lyne, A. G., Bailes., M., Kaspi, V. M., Qiao, G. and D'Amico, N., *Astrophys. J.*, 1992, **387**, L37–L41.
19. Kaspi, V. M., Johnston, S., Bell, J. F., Manchester, R. N., Bailes, M., Bessell, M., Lyne, A. G. and D'Amico, N., *Astrophys. J.*, 1994, **423**, L43–L45.
20. Bell, J F., Bessell, M. S., Stappers, B. W., Bailes, M. and Kaspi, V. M., *Astrophys. J.*, 1995, **447**, L117–L119.
21. Underhill, A. and Doazan, V., *B Stars With and Without Emission Lines.*, NASA Monograph Series, 1982.
22. Struve, O., *Astrophys. J.*, 1931, **73**, 94–103.
23. Waters, L. B. F. M., *Astr. Astrophys.*, 1986, **162**, 121.
24. Johnston, S., Lyne, A. G., Manchester, R. N., Kniffen, D. A., D'Amico, N., Lim, J. and Ashworth, M., *Mon. Not. R. Astron. Soc.*, 1992, **255**, 401–411.
25. Johnston, S. Manchester, R. N., Lyne, A. G. and D'Amico, N., *IAU Telegram* No. 5794, 1993.
26. Taylor, J. H. and Cordes, J. M., *Astophys. J.*, 1993, **411**, 674–684.

27. Johnston, S., Manchester, R. N., Lyne, A. G., Nicastro, L. and Spyromilio, J., *Mon. Not. R. Astron. Soc.*, 1994, **268**, 430–436.
28. Manchester, R. N. and Johnston, S., *Astrophys. J.*, 1995, **441**, L65–L68.
29. Watson, M and Bailes, M., unpublished, 1992.
30. Cominsky, L., Roberts, M. and Johnston, S., *Astrophys. J.*, 1994, **427**, 978–983.
31. Greiner, J., Tavani, M. and Belloni, T., *Astrophys. J.*, 1995, **441**, L43.
32. Kochanek, Ç. S., *Astrophys. J.*, 1993, **406**, 638–650.
33. King, A. and Cominsky, L., *Astrophys. J.*, 1994, **435**, 411–415.
34. Lipunov, V. M., Nazin, S. N., Osminkin, E. Y. and Prokhorov, M. E., *Astr. Astrophys.*, 1994, **282**, 61–72.
35. Tavani, M., Arons, J. and Kaspi, V. M., *Astrophys. J.*, 1994, **433**, L37–L40.
36. Manchester, R. N., Johnston, S., Lyne, A. G., D'Amico, N., Bailes, M. and Nicastro, N., *Astrophys. J.*, 1995, **445**, L137.
37. Illarionov, A. F. and Sunyaev, R. A., *Astr. Astrophys.*, 1975, **39**, 185–195.
38. Ghosh, P. and Lamb, F. K., *Astrophys. J.*, 1979, **234**, 296–316.
39. Ghosh, P., *Astrophys. J.*, 1995 (in press).
40. Johnston, S., Manchester, R. N., Lyne, A. G., D'Amico, N., Bailes, M., Gaensler, B. M. and Nicastro, L. *Mon. Not. R. Astron. Soc.*, 1995 (submitted).
41. Melatos, A., Johnston, S. and Melrose, D. B., *Mon. Not. R. Astron. Soc.*, 1995, **275**, 381–397.
42. Dougherty, S. M., Waters, L. B. F., Burki, G., Coté, J., Cramer, N., van Kerkwijk, M. H. and Taylor, A. R., *Astr. Astrophys.*, 1994, **290**, 609–622.
43. Kaspi, V. M., Tavani, M., Nagase, F., Hirayama, M., Hoshino, M., Aoki, T., Kawai, N. and Arons, J., *Astrophys. J.*, 1995 (submitted).
44. Tavani, M. and Arons, J., *Astrophys. J.*, 1995 (in preparation).
45. Grove, J. E., Tavani, M., Purcell, W. R., Johnson, W. N., Kurfess, J. D., Strickman, M. S. and Arons, J., *Astrophys. J.*, 1995, **447**, L113–115.
46. McCollum, B., Castelaz, M. W. and Bruhweiler, F. C., *Astrophys. J.*, 1995 (in preparation).
47. Unger, S. J. and Johnston, S. *Mon. Not. R. Astron. Soc.*, 1995 (in preparation).
48. Johnston, S., *et al.*, *Mon. Not. R. Astron. Soc.*, 1995 (in preparation).

Received 3 July 1995; revised accepted 12 July 1995.

A high-precision technique using X-ray reflectivity for the measurement of surface and interface roughness

P. V. Satyam, D. Bahr*, S. K. Ghose, G. Kuri, B. Sundaravel, B. Rout and B. N. Dev

Institute of Physics, Sachivalaya Marg, Bhubaneswar 751 005, India
*Hamburg Synchrotron Radiation Laboratory, HASYLAB at DESY, Notkestr. 85, D-22603 Hamburg, Germany

Surface and interface roughness plays a crucial role in ultrathin two-dimensional layers. We have set up an X-ray reflectometry facility for the measurement of surface and interface roughness with a precision better than 1 Å. As ion implantation is one of the techniques to fabricate ultrathin buried epitaxial layers, we have explored the effect of MeV ion implantation on surface roughness in two cases: Au-implanted $LiNbO_3$(001) and Co-implanted Si(111) samples. In both cases we have observed an implantation-induced enhancement of surface roughness. In addition, for the Co-implanted samples, we found unexpected interference fringes with dose-dependent periodicity.

HIGH-quality thin films are necessary for the fabrication of lower-dimensional structures for fundamental studies and for device fabrication in microelectronics. For the growth of high-quality films the role of solid–vacuum and solid–solid interfaces is very crucial. The quality of the thin film and the interfaces should be as good as possible. Therefore, one needs experimental probes to determine this quality. Among the relevant questions to be answered are the following: (i) What is the structure of the ultrathin grown layer, or what is the initial stage of growth? (ii) For a multielemental film, which type of atoms lie in the first atomic layer at the interface? (iii) What is the coordination number of the interface atoms? (iv) What is the distance between adjacent atomic layers across the interface? These questions are answered by using a very powerful technique involving generation of standing waves of X-rays[1–4]. One would also like to know the amount of statistical disorder, or roughness, on the film surface and on the buried film–substrate interface. This information along with the film thickness can be obtained from X-ray specular reflectivity measurements. Both the X-ray standing wave and the X-ray reflectivity measurement techniques are nondestructive and the high penetration power of X-rays enables the probing of buried layers and interfaces.

For the determination of interface roughness, the problem is to determine the correct probability density function for the interface position in the direction of the surface normal. This is mostly a Gaussian distribution. Then the spread in the interface position can be evaluated. X-ray reflectivity under the grazing incidence condition is a well-established technique to investigate this spread in the position of exposed (outer) and buried (inner) interfaces[5].

We have recently set up the grazing incidence X-ray reflectivity experiment for the measurement of surface and interface roughness in the Institute of Physics, Bhubaneswar. To our knowledge, this is the first high-precision grazing incidence X-ray reflectometry facility in India. Here we present a brief description of the set-up and some preliminary results.

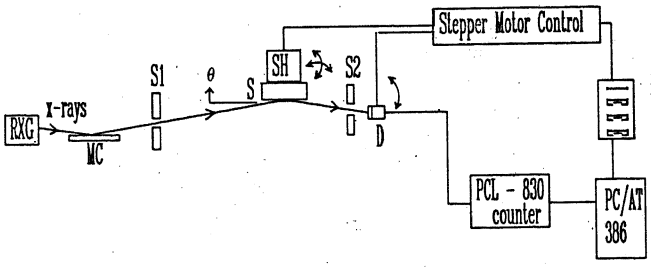
The experimental set-up uses an 18 kW rotating-anode X-ray source (RIGAKU 300VE) with a Mo anode. A line beam of Mo-$K_{\alpha 1}$ (17.479 keV) X-rays is selected with a symmetrically cut dislocation-free Si(111) monochromator crystal. The θ–2θ scan is performed using a HUBER four-circle diffractometer. The reflected X-rays are detected with a BaF_2 scintillation detector. The programs for the angular scans and data collection have been written in GWBASIC. A schematic diagram of the experimental set-up is shown in Figure 1. We also measure inelastic spectra, e.g. fluorescence, from the sample with a Si(Li) detector simultaneously with the reflectivity as a function of θ. The signal from the Si(Li), which first goes to a multichannel analyser, is automatically transferred to the computer (PC-386) memory along with the reflectivity data. The whole process is automated and controlled through the PC. This completes the set-up for X-ray standing wave experiment under grazing incidence specular reflection. This additional feature with the Si(Li) detector has not been shown in Figure 1. The details of the instrumentation aspect will be published elsewhere. A part of the instrumentation has already been published[6].

For X-rays the (real part of) index of refraction of any solid or liquid medium is less than unity. Thus, it is possible to obtain total external reflection from matter. For the two-media system the reflectivity is obtained from the Fresnel coefficient of reflection[5]:

$$r_{1,2} = \frac{k_{z,1} - k_{z,2}}{k_{z,1} + k_{z,2}}, \tag{1}$$

where

$$k_{z,l} \equiv k_1(\theta_1^2 - 2\delta_l - 2i\beta_l)^{1/2}, \quad l = 1, 2, \tag{2}$$



RXG : Rotating anode X-ray Generator, MC : Monochromator,
S1,S2 : Slits, S : Sample, SH : Sample Holder, D : Detector,

Figure 1. The schematic experimental set-up for X-ray reflectivity measurements.

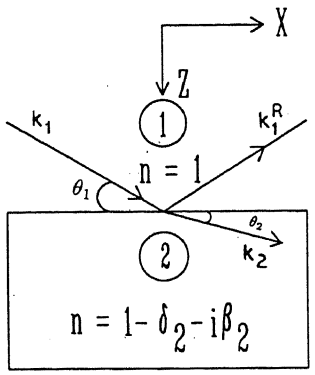


Figure 2. Sketch of X-ray reflection and refraction for a two-medium system.

θ_1 and k_1 being the glancing angle and the wave vector in medium 1, respectively (Figure 2);

$$n_l = 1 - \delta_l - i\beta_l \tag{3}$$

is the complex refractive index of the material of the lth layer. δ and β values are very small, each of the order of 10^{-5} or less. δ and β values are given by $\delta = r_e \lambda^2 N/2\pi$ and $\beta = \lambda\mu/4\pi$, where r_e is the classical electron radius, λ is the wavelength of the incident radiation, N is the total number of electrons per unit volume and μ is the linear absorption coefficient[7]. Equation (1) is valid for the s-polarized radiation. For the p-polarized radiation

$$r_{1,2}^{p} = \frac{k_{z,1}/n_1^2 - k_{z,2}/n_2^2}{k_{z,1}/n_1^2 + k_{z,2}/n_2^2}. \tag{4}$$

However, for X-rays ($n_l \approx 1$) eq. (4) reduces to eq. (1).

If the first medium is air ($\delta = \beta = 0$) and the second a solid substrate, the reflectivity is

$$R_{1,2} = |r_{1,2}|^2 = \left| \frac{\theta - (\theta^2 - 2\delta_2 - 2i\beta_2)^{1/2}}{\theta + (\theta^2 - 2\delta_2 - 2i\beta_2)^{1/2}} \right|^2. \tag{5}$$

(We have dropped the subscript of θ_1.) This is applicable for a smooth surface. (A surface is also called an air–solid or vacuum–solid interface.)

For a rough surface, the surface position can be described by a Gaussian probability distribution as introduced by Nevot and Croce[8]:

$$P(z) = \frac{1}{\sqrt{2\pi}\sigma_2} e^{-z^2/2\sigma_2^2}. \tag{6}$$

With this modification the Fresnel coefficient of reflection for a rough surface becomes

$$r_{1,2}^{\text{rough}} = r_{1,2} \exp^{-2\sigma_2^2 k_{z,1} k_{z,2}} \tag{7}$$

in the limit $\sigma_2 k_{z,2} < 1$; σ_2 is the root mean square roughness of the surface of medium 2, and $r_{1,2}$ is given by eq. (1). This result was also obtained by Sinha *et al.*[9] in a different approach within the distorted-wave Born approximation (DWBA). Thus, the reflectivity (eq.(5)) gets modified for a rough surface:

$$R_{1,2}^{\text{rough}} = \left| r_{1,2}^{\text{rough}} \right|^2. \tag{8}$$

Here we present some results of reflectivity measurements on Au-implanted $LiNbO_3$(0 0 1) and Co-implanted Si(1 1 1) single-crystal substrates. $LiNbO_3$ is a widely used dielectric material for optical waveguides. The waveguiding properties are controlled by incorporating impurities, e.g., Ti, in the $LiNbO_3$ lattice. It is important to determine the impurity site in the lattice[10]. For the lattice location of Au in $LiNbO_3$ we performed X-ray standing wave and ion-channeling experiments. These results will be published elsewhere. Here we present the results of measurements of surface roughness of an implanted sample.

Interest in Co implantation in Si arises from the fact that it is possible to fabricate a buried epitaxial $CoSi_2$ layer in Si by ion implantation of Co followed by annealing. This process is known as ion beam synthesis (IBS). The IBS-prepared silicides have received much attention as their electrical transport properties are better than those prepared by the conventional ultrahigh vacuum procedure. A possible application of an IBS-fabricated Si/$CoSi_2$/Si heterostructure is as a metal base transistor[11]. Measurement and control of interface roughness in these structures are important. The results of surface roughness measurements on Co-implanted Si(1 1 1) samples are presented here and compared with those on the virgin Si(1 1 1) sample.

Figure 3 shows the X-ray specular reflectivity from a virgin (0 0 1)-oriented $LiNbO_3$ sample as a function of the grazing incidence angle. The curve fitted to the data corresponds to a surface roughness of 9.6 Å. In order to elucidate the precision of the roughness measurement, the theoretical curves for two values of roughness of 8.6 and 10.6 Å are also shown. These curves are clearly different from the best fit to the data, implying that the precision is better than 1 Å. The theoretical curves obtained from eq. (8) have been convoluted with the beam divergence (0.05°) for fitting to the experimental data. Results from the measurement on Au-implanted sample are shown in Figure 4, which also shows the virgin $LiNbO_3$ result for comparison. The measured roughness on the implanted sample is 14.5 Å, i.e. the surface roughness has been enhanced due to implantation. (It may be noted that high-energy ion implantation can produce interesting viscoelastic effect, leading to surface smoothing or roughening[12].) We have not made measurements on the implanted sample before annealing.

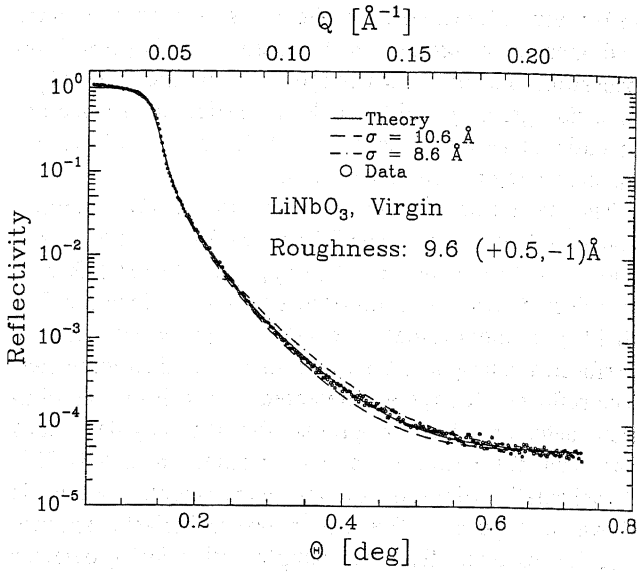


Figure 3. Reflectivity as a function of the grazing angle of incidence θ for a virgin $LiNbO_3$(0 0 1) sample. The value of the momentum transfer $Q = 4\pi \sin\theta/\lambda$ is shown on the top abscissa.

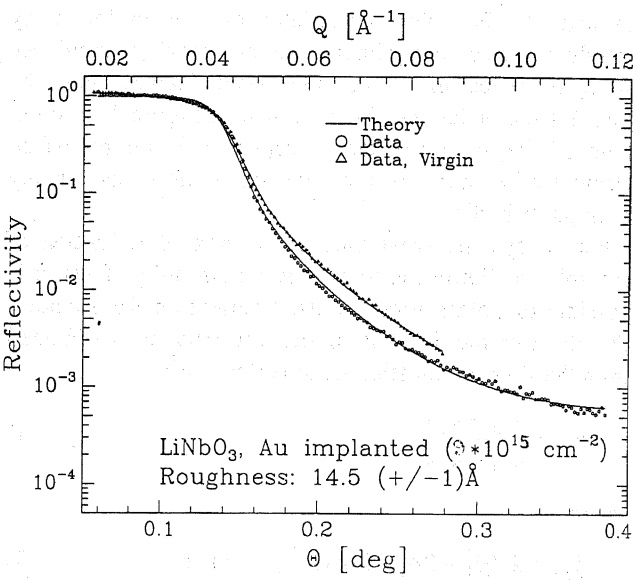


Figure 4. Reflectivity from an Au-implanted $LiNbO_3$(0 0 1) sample (O). For comparison the reflectivity from a virgin sample is also shown.

Surface roughness might have been even higher on the as-implanted sample. Au implantation was done using the 3 MV Pelletron accelerator in our Institute. 1.38 MeV Au^{2+} ions were implanted into the $LiNbO_3$ sample at a tilt angle (the angle between the surface normal and the ion beam direction) of 65°. The sample was annealed at 450°C for 40 min under flowing oxygen to remove implantation-induced lattice damage. The ion-channelling measurement showed that the bulk lattice damage was almost fully removed.

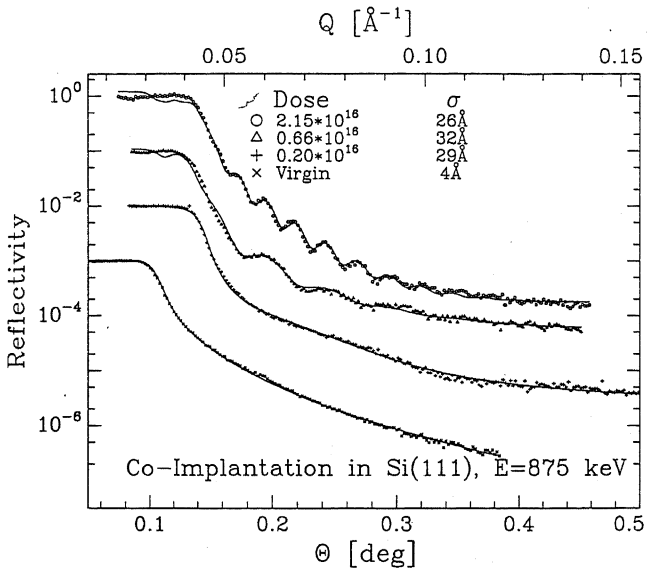


Figure 5. Reflectivity from Co-implanted Si(1 1 1) samples with various Co doses (ions/cm^2) and from a virgin Si(1 1 1) sample is shown. The surface roughness values (σ) are listed. The ordinates of the successive curves have been shifted by a decade for clarity.

Co^+ implantation was done at a tilt angle of 70°. The samples were not annealed. The reflectivity results are shown in Figure 5. A comparison with the virgin Si(1 1 1) sample shows that the surface roughness is significantly enhanced. However, the roughness is not strongly dose-dependent within the range of doses (2×10^{15}–2.15×10^{16} Co/cm^2) studied. The oscillations on the reflectivity curves arise from the interference due to scattering from a carbon layer deposited on the surface during implantation. The thickness of this layer is determined from the periodicity of these oscillations. Independently, we have determined the thickness of this carbon film and the depth distribution of Co atoms by Rutherford backscattering spectrometry[13]. Detailed analysis of the Co-implanted system will be published elsewhere.

X-ray reflectivity from a multilayer system has been worked out by Parrat[5]. For the multilayer system each interface can be represented by a Gaussian of the form given in eq. (6), with σ_l representing the rms roughness of the $(l-1)$th interface. (The first layer being vacuum, σ_2 represents the roughness of the surface of the second layer or the first interface.)

We believe that now with the availability of this powerful technique in India, further activities will be stimulated in the fast-developing area of surface and interface physics.

1. Dev, B. N., Materlik, G., Grey, F., Johnson, R. L. and Calusnitzer, M., *Phys. Rev. Lett.*, 1986, **57**, 3058–3061.
2. Vlieg, E., Fischer, A. E. M. J., van der Veen, J. F., Dev, B. N. and Materlik, G., *Surf. Sci.*, 1986, **178**, 36–46; Zegenhagen, J., Huang, K.-G., Gibson, W. M., Hunt, B. D. and Schowalter, L. J., *Phys. Rev.*, 1989, **B39**, 10254–10260.
3. Fischer, A. E. M. J., Vlieg, E., van der Veen, J. F., Clausnitzer, M. and Materlik, G., *Phys. Rev.*, 1987, **B36**, 4769–4773.
4. Satyam, P.V. and Dev, B. N., *Indian J. Phys.*, 1994, **A68**, 23–32.
5. Parrat, L. G., *Phys. Rev.*, 1954, **95**, 359–369.
6. Satyam, P. V. and Dev, B. N., *Indian J. Phys.*, 1995, **A69**, 417–427.
7. James, R. W., *The Optical Principles of the Diffraction of X-rays*, Ox Bow, Woodridge, 1982, Chap. 13.
8. Nevot, L and Croce, P., *Rev. Phys. Appl.*, 1980, **15**, 761.
9. Sinha, S. K., Sirota, E. B., Garoff, S. and Stanley, H. B., *Phys. Rev.*, 1988, **B38**, 2297–2311.
10. Gog, Th., Harasimowicz, T., Dev, B. N. and Materlik, G., *Europhys. Lett.*, 1994, **25**, 253–258.
11. Vanderstrten, H., Bruynseraede, Y., Vantomme, M. F., Wu, A., Langouche, G. and Phillips, J. M., *Appl. Phys. Lett.*, 1990, **57**, 135–137.
12. Gutzmann, A., Klaumuenzer, S. and Meier, P., *Phys. Rev. Lett.*, 1995, **74**, 2256–2259.
13. Satyam, P. V., Sundaravel, B., Ghose, S. K., Rout, B., Bahr, D. and Dev, B. N. (to be published).

ACKNOWLEDGEMENTS. Prof. G. Materlik of Hamburg Synchrontron Radiation Laboratory (HASY-LAB at DESY) provided us the monochromator crystal and a slit system. The short visit of Dr D. Bahr from Prof. Materlik's group was supported by our Indo-German bilateral collaboration (Physics-25). We thank Prof. V. S. Ramamurthy for his support and encouragement.

Received 4 May 1995; revised accepted 20 July 1995

Single-ion activities by a solid ion transmitter bridge and a reference electrode without liquid junction

S. Parthasarathy and K. Ramya

Centre for Electrochemical and Energy Research, SPIC Science Foundation, 110, Mount Road, Madras 600 032, India

It had been shown that it is possible to measure thermodynamic values of single-ion activity using a solid ion conductor to bridge a reference electrode and a test electrode, thus eliminating liquid junction potentials. This led to redefinition of activity of an ion and brought out into sharp focus the crucial role of the charge factor in determining the effective strength of an ion. In this paper, we have measured potentials, free from liquid junction, using chloride ion sensors and the solid-ion transmitter bridge for determining chloride ion activity in zinc chloride solutions, leading to confirmation of definition of activity of an ion already proposed.

THE central paradigm of the theory of ionic solutions of strong electrolytes can be stated as follows:

1. It is impossible to measure the activity of single ion in solutions of strong electrolytes, as any reference

electrode will have to necessarily form a liquid junction with the test solution[1,2].

2. Liquid junctions potentials cannot be eliminated, but can at best be minimized using devices like a salt bridge[3,4].

3. The concept of individual activity coefficients cannot be defined accurately, and such coefficients may not even be determined experimentally without some supplementary definition of nonthermodynamic nature[5].

Harned[6] has expressed the point at issue in explicit terms: 'We are confronted with interesting perplexity that it is not possible to compute liquid junction potentials without a knowledge of individual ion activities, and it is not possible to determine individual ion activities without an exact knowledge of liquid junction potentials. For the solution of this difficult problem, it is necessary to go outside the domain of exact thermodynamics'. Correlating the liquid junction potentials and individual ionic activities, Guggenheim[7] has deduced equations for cell potentials with liquid junction that are free from ionic activities and concluded that the activity coefficient of a single ionic species is not an operational concept and is completely unnecessary for an adequate treatment of the thermodynamics of the cell with liquid junction. Conversely, this type of cell can furnish no exact information regarding the activity of a single ionic species, whereas the meaning and usefulness of single ion activities have been stressed by Bronsted *et al.*[8], MacInnes[9] and Kortum[10].

Single-ion activity (SIA) would be a prerequisite for building electrochemical potential series without invoking arbitrary zero potential for the standard hydrogen electrode. Measurement of SIA would provide a tool for direct measurement of free metal ion concentration for studies on metal–ligand interactions instead of the fitment procedure adopted currently, using only ionic strength and pH, by successive iterations to calculate the activity of the species in solution.

Thermodynamic activities of the electrolyte can be obtained from the potential of an electrochemical cell without liquid junction. For example, in the cell Zn, $ZnCl_2$/AgCl, Ag system consisting of Zn electrode and an indicator electrode, Ag/AgCl in the same solution is treated as a cell without liquid junction and thermodynamic activities are derived from cell potentials.

The activity of an ionic species is arbitrarily defined as the product of molality and activity coefficient. If SIA can be determined experimentally, this arbitrary assumption can be dispensed with, by comparing it with value calculated using molality and the single-ion activity coefficient (calculated using Debye–Huckel equation and its modifications)[11–13]. This would enable us to check whether the present mode of calculation of SIA will require any modification about which reservations have already been expressed[14].

It has already been shown[15] that the activity of an ion in a solution of molality m may be redefined as the effective strength of the ion expressed by the term $Z^2_i m\gamma_i$ (where Z_i is the valence of the ion), rather than the effective concentration signified by $m\gamma_i$. EMF measurements of the cell Zn/$ZnCl_2$(m)/bridge/KCl(satd), AgCl/Ag have been carried out for different concentrations (molality) of zinc chloride, using a bridge based on solid silver chloride ionic conductor to avoid liquid junction potential. From the experimental data the activity of zinc ion has been derived for different molal concentrations of zinc chloride and has been found to compare well with the calculated values.

This paper presents a new approach to eliminate liquid junction potential by using a solid ion transmitter bridge for determining single-ion activities (chloride ion) in zinc chloride solutions. The experimental results lend further support to the new definition of activity of an ion proposed earlier[15]. The preparation and other details of the solid ionic conductor bridge are also discussed in this paper.

The chloride-ion-selective electrodes are the Ag/AgCl electrodes prepared by dipping a loop of high-purity silver wire into molten high-purity silver chloride. These electrodes are then tested against saturated calomel electrode (SCE) for reproducibility and stability in the cell Ag/AgCl, KCl(satd), Hg_2Cl_2/Hg. Electrodes with a microcapillary tube supplied by Tacussel Electronique, France, were found to be most suitable. The potential is independent of the activity of chloride ions as the cell reaction

$$Ag_2Cl_2 + 2Ag \rightleftharpoons 2Hg + 2AgCl$$

involves only solid substances, and is found to be 0.0456 V (ref. 16). The electrodes were found to attain equilibrium potential faster than thermal or thermal electrolytic electrodes[17]. The zinc electrodes were specpure zinc rods supplied by Johnson Matthey, UK, sleeved by glass tube, fixed with epoxy with 1 cm length of rod exposed, which is amalgamated with distilled mercury. The procedure for preparation of zinc chloride solutions and the experimental set-up used is the same as described elsewhere[15].

Solid AgCl was used to make solid ion transmitter because of its unique plasticity characteristics. For this purpose, pure AgCl was melted into a bridge in the form of a 'U' tube made from pyrex glass with a capillary of 2 mm bore diameter so that it is a uniform solid on par with rolled or cast AgCl sheet or rod. The tubes are put in the furnace maintained at 380°C and held at this temperature for 10 h and then allowed to cool slowly. This will eliminate porosity, cracks and fissure formation and maintain continuity of the solid chloride between the two tips of the bridge.

The bridges were subjected to:

1. *Pressure test.* On application of gas at a high pressure from a nitrogen cylinder, there was no pressure drop, thereby confirming that there are no crevices, leaks or cracks in the solid AgCl.

2. *Electrochemical test.* The measured potential of the cell Ag/AgCl, KCl(satd)/Bridge/KCl(satd), Hg_2Cl_2/Hg was 45.6 mV in accord with expected value.

3. *Creep test.* One arm of the bridge was placed in saturated KCl solution and the other in $AgNO_3$ solution. No turbidity or precipitate of AgCl was found even after 24 h. It was also found that the potential-measuring capability is not affected when we switch over from saturated KCl to $ZnCl_2$ (0.1 m) solution as shown in the experiment. These tests confirm that there is no creep of Cl^- ions from one cell to another.

A potential of 1000 V was applied across the two ends of the bridge by connecting to a DC power source and at a point 1 cm away from tips by fusing a Pt wire. The current was in the picoampere range and there was no precipitate of AgCl formed when the bridge was dipped in $AgNO_3$ solution; the potential was applied for several hours. No evolution of gas (chlorine or hydrogen) was observed, thus ruling out the possibility of corrosion reactions. The conductivity of AgCl is 3×10^{-8} mho/cm.

It is noteworthy that the potentials do not vary irrespective of whether a thin bent sheet of rolled AgCl or a glass tube filled with cooled molten AgCl is used. Therefore, the question of creep of KCl through capillary crack or glass tube will not arise. Solid AgCl bridge cannot, therefore, be considered as an equitransferent bridge like KCl.

The activity of zinc chloride electrolyte was determined by measuring the potential of the cell without liquid junction between the Zn electrode and a Ag/AgCl electrode in $ZnCl_2$ solution, i.e. the cell $Zn/ZnCl_2(m)/AgCl/Ag$.

The activity of the chloride ion was obtained by measuring the potential of chloride-ion-responsive electrode in the respective zinc chloride solutions against the Ag/AgCl reference electrode of fixed potential of 0.199 V (ref. 18), i.e., the potential of the cell

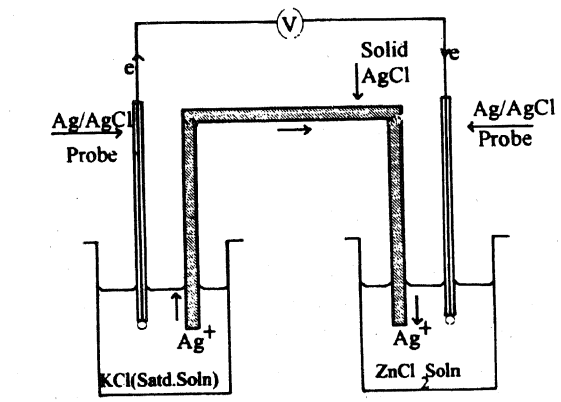
Ag/AgCl, KCl(satd)/Bridge/$ZnCl_2$(m), AgCl/Ag.

AgCl has high ohmic resistance; and it is, therefore, necessary to use a voltage-measuring instrument of very high input impedence for handling singals from high-resistance sources, with negligible IR drop. The instrument used for this study was ORION expandable ion analyser, model no. EA 940, from ORION Research Inc., USA.

It is more convenient to measure the difference in potentials between the various $ZnCl_2$ solutions as the impedences of the two solutions are closer. The experiments carried out for this purpose involve measurement of potentials of the following cells:

1. Ag/AgCl, $ZnCl_2$ 0.1 m/bridge/$ZnCl_2$ 0.01 m, AgCl, Ag.
2. Ag/AgCl, $ZnCl_2$ 0.1 m/bridge/$ZnCl_2$ 0.001 m, AgCl, Ag.

The AgCl-based bridge is best described as a solid-state ion conductor. If the current in the picoampere range is carried entirely by Ag^+ ions, the schematic will be as in Figure 1 *a*, which is in keeping with the normal assumption that Ag^+ ion is the migrant ion in solid AgCl[19,20]. If chloride is to be considered, the reaction schematic can be visualized as in Figure 1 *b*. In any case both the schemes are equivalent. The actual studies on the electromigration of ions is beyond the scope of this paper. But the fact remains that without liquid junction the bridge is able to establish contact between the reference and the test solutions. Further, the instrument used for



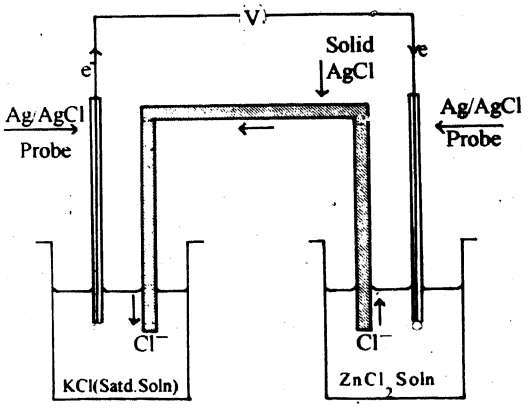
$Ag + Cl^- \rightarrow AgCl + e^-$ (Probe)

$AgCl \rightarrow Ag^+ + Cl^-$ (Bridge)

$e^- + AgCl \rightarrow Ag + Cl^-$ (Probe)

$Ag^+ + Cl^- \rightarrow AgCl$ (Solid) (Bridge)

b



$Cl^- + Ag \rightarrow AgCl + e^-$ (Probe)

$e^- + AgCl \rightarrow Ag + Cl^-$ (Probe)

Figure 1. Schematic for: *a*, silver ion transmission; *b*, chloride ion transmission.

Table 1. Activities of ions in $ZnCl_2$ solutions with related data at 25°C

Description	Symbol	Data			Remarks
Molality	m	0.1	0.01	0.001	
Activity coefficient					
Chloride ion	γ_{Cl^-}	0.709	0.841	0.938	
Potentials (V)					
$Zn/ZnCl_2/AgCl/Ag$	e_1	1.082	1.157	1.237	Observed values
$Ag/AgCl$, $KCl(satd)/ZnCl_2$, $AgCl/Ag$	e_2	0.0915	0.1463	0.2024	Observed values
$Ag/AgCl$, Cl^- electrode	e_3	0.2905	0.3453	0.4014	$[e_2 + 0.1990]$
Activities					
Observed	a_{Cl^-}	0.0706	0.00847	0.00096	$\text{antilog}\frac{(0.2224 - e_3)}{0.05916}$
Calculated	a_{Cl^-}	0.0709	0.00841	0.00094	$m\gamma_-$
By textbook method	a_{Cl^-}	0.1418	0.01682	0.00188	$2m\gamma_-$
Calculated	$a_{Zn^{2+}}$	0.0998	0.0207	0.0033	$a_{ZnCl_2}/a^2_{Cl^-}$
Observed	$a_{Zn^{2+}}$	0.0998	0.0212	0.0032	ref. 14
By textbook method	$a_{Zn^{2+}}$	0.0252	0.005	0.00077	$a_{ZnCl_2}/a^2_{Cl^-}$

$E_{Ag/AgCl}$, KCl(satd) = 0.199 V; $E^0_{Ag/AgCl} = 0.2224$ V.

measuring the potentials has an input impedence of $10^{12}\ \Omega$. Being a solid-state device and the mechanism clearly a Grotthuss type of conduction of the ion inside the bridge, the requirement that the transference number of conducting ion should be unity will not arise.

A membrane potential cannot be associated with AgCl, as it can by no means function as a membrane. AgCl, being ionized in the solid state, will not permit absorption of either negative or positive ions.

Table 1 presents the potentials measured to calculate the activity of the electrolyte and the chloride ion.

For the cell $Zn/ZnCl_2(m)/AgCl/Ag$, the reactions are:

$$Zn \rightarrow Zn^{2+} + 2e^- \quad (1)$$
$$2AgCl + 2e^- \rightarrow 2Ag + 2Cl^- \quad (2)$$
$$Zn + 2AgCl \rightarrow Zn^{2+} + 2Cl^- + 2Ag$$

Following the IUPAC convention, the cell potential e_1 is given by

$$e_1 = E^0_{Ag/AgCl} - E^0_{Zn/Zn^{2+}} - (RT/2F)\ln[a_{ZnCl_2}]. \quad (3)$$

The electrolyte activity for different concentrations calculated from the e_1 values are given in Table 2. The activity values are in close agreement with those reported earlier[21–23].

For measuring the chloride ion activity, a liquid-junction free solid ion transmitter was used to bridge the reference half cell and the half cell comprising the test solutions. The cell set-up used was:

$Ag/AgCl$, $KCl(satd)/bridge/ZnCl_2(m)$, $AgCl/Ag$.

As the liquid junction potential is zero, following IUPAC convention, the potential e_2 of the cell is given by

$$e_2 = E^0_{Ag/AgCl} + (RT/F)\ln(1/a_{Cl^-}) - E_{ref}.$$

If we denote, $E_2 + E_{ref}$ as e_3,

$$a_{Cl^-} = \text{antilog}\,[(0.2224 - e_3)/0.05916].$$

Equation (3) can be split as

$$E_1 = E^0_{Ag/AgCl} + (RT/2F)\ln 1/a^2_{Cl^-}), \quad (4)$$

$$E_1 = E^0_{Ag/AgCl} - (RT/2F)\ln a_{Cl^-}, \quad (5)$$

$$E_2 = E^0_{Zn/Zn^{2+}} + (RT/2F)\ln a_{Zn^{2+}}. \quad (6)$$

In a completely dissociated strong electrolyte system like MA_2, e.g. $ZnCl_2$, there is an inherent asymmetry arising out of the charge factor, the cation being doubly charged and the anion being singly charged. It is this asymmetry that is reflected in the Nernst equation appropriate to the anion, which can reflect only the activity of each dissociating fragment. In terms of the electrolyte molality m, the molality of cation is m and that of the anion will be $2m$. Then the question arises as to whether which of these two factors 'charge' or 'total molality' is relevant to arrive at the ionic activity. If the charge factor is ignored in Cl^- ion, the activity of the chloride ion will be $2m\gamma_-$ and will contribute a $(2m\gamma)^2$ term in the expression for the activity of the electrolyte. On the other hand, if charge is the factor to be reckoned with, the chloride ion will contribute only $(m\gamma_-)^2$ term to the expression for the activity of the electrolyte.

Table 2. Activities of $ZnCl_2$ electrolyte

Concentration (m)	0.1	0.01	0.001
Activity	5.017×10^{-4}	1.462×10^{-6}	2.89×10^{-9}

Table 3. Difference in potential measured against reference solutions at 25°C

Description	Symbol	Data		Remarks
Molality	m	0.01	0.001	
Cell potential (V)				
Ag/AgCl, $ZnCl_2$, 0.1/$ZnCl_2$(m), AgCl/Ag	e_4	0.055	0.111	
Ag/AgCl, Cl^- electrode	e_5	0.3451	0.4011	$e_4 + 0.2901$

Table 4. Calculation of electrolyte activity

Currently used method	New method proposed
$a_{ZnCl_2} = a_{Zn^{2+}} \times a_{Cl^-} \times a_{Cl^-}$	$a_{ZnCl_2} = a_{Zn^{2+}} \times a_{Cl^-} \times a_{Cl^-}$
$= m\gamma_+ \times 2m\gamma_- \times 2m\gamma_-$	$= 4m\gamma_+ \times m\gamma_- \times m\gamma_-$
$= 4m^3\gamma_\pm^3$	$= 4m^3\gamma_\pm^3$

Clearly, the current definition of activity of an ion as a product of its concentration m and the activity coefficient needs to be revised as the experiments show that the activity is the same as the effective ionic strength defined as $(Z_i^2 m)\gamma_i$ (ref. 15), where Z_i is the valency of the ion.

Table 3 represents the difference in potentials between the chloride-ion-responsive probes in different $ZnCl_2$ solutions. Ag/AgCl, $ZnCl_2$ (0.1) has been used as reference solution with an electrode potential of 0.2901 V (ref. 15). These results confirm the activity of the chloride ion obtained using Ag/AgCl electrode in saturated KCl as the reference electrode shown in Table 1.

The activity of the zinc ion in $ZnCl_2$ solution can be obtained from the activity of the electrolyte and activity of chloride ions, i.e.

$$a_{Zn^{2+}} = a_{ZnCl_2} / a_{Cl^-}^2 .$$

This is found to be in close agreement with our earlier measurement[15] of activity of zinc ion using solid ion transmitter bridge. In the case of zinc ion in $ZnCl_2$ solutions, the charge factor assumes greater importance, and, if activity is treated as the product of the ionic strength of the zinc ion and its activity coefficient, the expected value for the zinc ion will be $4m\gamma_+$. The results of measurements of zinc ion activity using solid ion transmitter is in agreement with the value of $4m\gamma_+$.

The activity of the electrolyte calculated using the current definition to activity of ion and the new definition proposed leads to $a_{ZnCl_2} = 4m^3\gamma_\pm^3$. This is just an accidental coincidence and would not be the case with all electrolytes.

Bates[14] in his recent review on single-ion activities covering a period of 75 years has concluded that individual ion activities are not amenable to thermodynamic measurements. However, in this paper we have shown that liquid junction potentials can be eliminated, thus paving the way for thermodynamic analysis of single-ion activity, overcoming the impasse highlighted by Harned[6].

1. Glasstone, S., An *Introduction to Electrochemistry*, Affiliated East West Press, New Delhi, 1975, pp. 230–231.
2. Crow, D. R., *Principles and Applications of Electrochemistry*, Chapman and Hall, London, 1988, pp. 6–7.
3. Bard, A. J. and Faulkner, L. R., *Electrochemical Methods – Fundamentals and Applications*, Wiley, New York, 1980, p. 72.
4. Atkins, P. W., *Physical Chemistry*, English Language Book Society, Oxford University Press, Oxford, 1987, p. 348.
5. Guggenheim, E. A., *J. Phys. Chem.*, 1929, **33**, 842–849.
6. Taylor, H. S., *A Treatise on Physical Chemistry*, Van Nonstrand, New York, 1924, vol. II, p. 782.
7. Guggenheim, E. A., *J. Phys. Chem.*, 1930, **34**, 1758–1766.
8. Bronsted, J. N., Delbanco, A. and Bolquartz, K., *Z. Phys. Chem.*, 1932, **A162**, 128–146.
9. MacInnes, D. A., *The Principles of Electrochemistry*, Reinhold, New York, 1939, Chap. 13.
10. Kortum, G., *Electrolytlösungen*, Becker and Eris Kom Ges., Leipzig, 1941, Chap. 12.
11. Debye, P. and Huckel, E., *Phys. Z.*, 1923, **24**, 185–206.
12. Guntelberg, E., *Z. Phys-Chem.*, 1926, **123**, 199–247.
13. Davies, C. W., *J. Chem. Soc.*, 1938, 2093–2098.
14. Bates, R.G., *Am. Chem. Soc. Symp. Ser.*, 1989, **390**, 142–152.
15. Parthasarathy, S. and Srinivasan, K. R., *Curr. Sci.*, 1994, **66**, 56–59.
16. Kortum, G., *Treatise on Electrochemistry*, Elsevier Publishing Co., London, 1965, p. 308.
17. Ives, D. J. G. and Janz, G. J., *Reference Electrodes – Theory and Practice*, Academic Press, New York, 1961, pp. 207–210.
18. Jeffery, J. H., Basset J., Mendham, J., Denney, R. C., *Textbook of Quantitative Analysis*, English Language Book Society/Longman, U.K., 1989, pp. 554.
19. Compton, W. D., *Phys. Rev.*, 1956, **101**,1209–1210.
20. Tannhauser, D., *Phys. Chem. Solids*, 1958, **5**, 224–235.
21. Scatchard, G. and Tefft, R. F., *J. Am. Chem. Soc.*, 1930, **52**, 2272–2281.
22. Lutffullah, Dunsmore, H. S. and Paterson, R., *J. Chem. Soc., Faraday Trans.*, 1976, **72**, 495–503.
23. Rohinson, R. A. and Stokes, R. H., *Trans. Faraday Soc.*, 1940, **36**, 740–748.

Received 17 May 1995; revised accepted 10 July 1995

Fall of liquid drops through a viscous medium: A fresh expression for drag force by the method of dimensions

G. J. Srinivasan and P. Satyanarayana
Forensic Sciences Department, Mylapore, Madras 600 004, India

An attempt has been made afresh by the method of dimensions to get a fresh expression for the drag force F arising from the motion of a liquid drop not exhibiting oscillation through a viscous medium, assuming that the drag force F is a function of D, u, σ, η and ρ alone (D, diameter of the drop; u, terminal velocity acquired by the drop; σ, density of the liquid in the column; η, viscosity of the column liquid; ρ, density of the liquid drop). This paper also relates to the experimental verification of the fresh drag force expression obtained in the case of a liquid drop with fairly high Reynolds number (20–305), Eotvos number (2×10^{-2} to 105×10^{-2}) and Morton number (4.75×10^{-12} to 3.75×10^{-7}) falling without oscillation through an immiscible liquid medium. Furthermore, the experimental value of the constant in the fresh drag force expression presented in this paper is found to be 3.6094.

SEVERAL authors have conducted studies on the fall of solid spheres[1–5] and liquid drops[6–30] through a viscous medium. So far, to the authors' knowledge, none has dealt with the problem of fall of liquid drops with only six variables, F, D, u, σ, η, ρ, in the dimensional analysis of drag force F. Therefore, an attempt has been made afresh to get a fresh expression for the drag force when the drops fall freely without oscillation with these six variables by dimensional analysis. Furthermore, an attempt has also been made with falling drops of fairly high Reynolds, Eotvos and Morton numbers to determine the value which the fresh drag force expression presented in this paper predicts for the constant K by experiment.

In general, there are at least six dimensionless groups that govern the motion of a drop, namely, $\pi_1 = F/D^2u^2\sigma$, $\pi_2 = \eta/\sigma uD$, $\pi_3 = \rho/\sigma$, $\pi_4 = \eta/\eta_0$, $\pi_5 = \gamma/Du^2\sigma$, $\pi_6 = u^2/gD$, where γ is the interfacial tension, D is the equivalent diameter of the drop and g is the acceleration due to gravity. η_0 is the viscosity of the drop. In general, the shape of the drops need not be spherical. The experimental work shows that for sufficiently small drops which appear to be ellipsoidal in shape over the range of Reynolds number $Re = 20$–305 ($Re = \sigma uD/\eta$), Eotvos number $Et = 2 \times 10^{-2}$–105×10^{-2} ($Et = g(\rho - \sigma)D^2/\gamma$), and Morton number $Mo = 4.75 \times 10^{-12}$ to 3.75×10^{-7} ($Mo = g\eta^4(\rho - \sigma)/\sigma^2\gamma^3$) the following simplification holds good:

$$\pi_1 = \phi(\pi_2, \pi_3) = K(\pi_2, \pi_3)^{1/2}, \qquad (1)$$

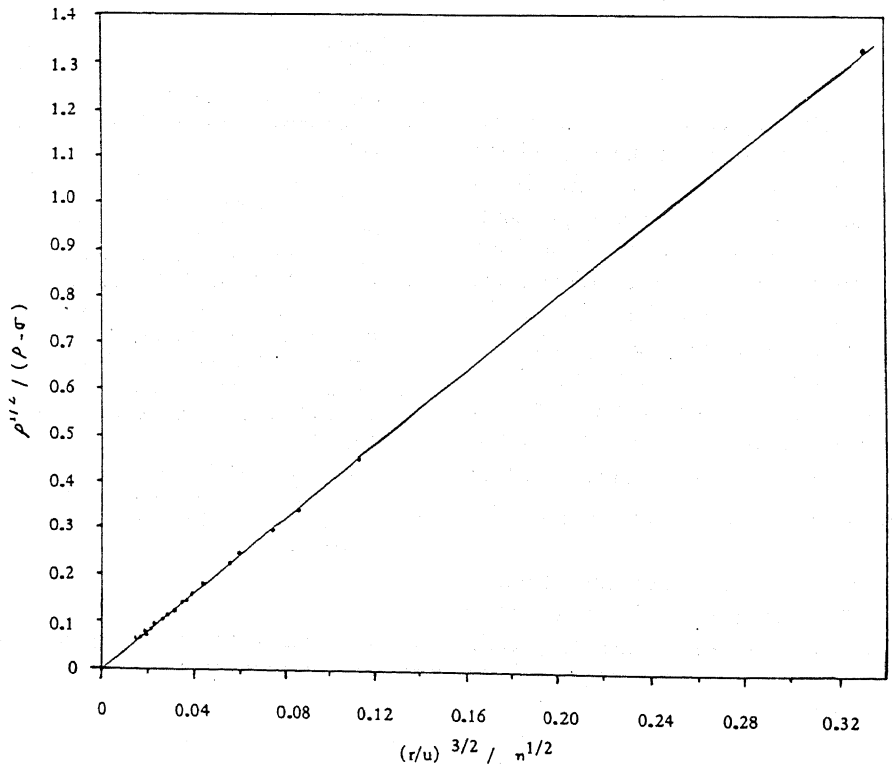


Figure 1. Plot of data given in Table 3.

Table 1. The systems and their density, ratio of density, viscosity, interfacial tension and ratio of radius to terminal velocity values

Liquid drop	Column liquid	ρ (kg m^{-3})	σ (kg m^{-3})	$\rho - \sigma$ (kg m^{-3})	ρ/σ	Column liquid η (N s m^{-2}) viscosity	Interfacial tension γ ($\times 10^{-3}$ N m^{-1})	r/u (s)
Eugenol	Water	1072.81	1000.00	72.81	1.07281	0.00100	12.82	0.0235
Chlorobenzene	Water	1097.99	1000.00	97.99	1.09799	0.00100	46.03	0.0195
Benzoyl benzoate	Water	1112.97	1000.00	112.97	1.11297	0.00100	34.59	0.0178
1-2-Dichloroethane	Water	1242.91	1000.00	242.91	1.24291	0.00100	31.78	0.0111
1-2-Dichlorobenzene	Water	1291.54	1000.00	291.54	1.29154	0.00100	84.77	0.0100
Trichloroethylene	Water	1444 05	1000.00	444.05	1.44405	0.00100	43.42	0.0077
Chloroform	Water	1477.40	1000.00	477.40	1.47740	0.00100	41.70	0.0073
Bromobenzene	Water	1492.21	1000.00	492.21	1.49221	0.00100	70.20	0.0071
Carbon tetrachloride	Water	1579.63	1000.00	579.63	1.57963	0.00100	39.06	0.0065
Nitrobenzene	Water	1193.92	1000.00	193.92	1.19392	0.00100	46.01	0.0125
Carbon disulphide	Water	1251.70	1000.00	251.70	1.25170	0.00100	31.36	0.0107
Dichloromethane	Water	1314.32	1000.00	314.32	1.31432	0.00100	36.71	0.0094
Water	Kerosene	1000.00	797.34	202.66	1.25417	0.00129	43.78	0.0125
Alcohol	Kerosene	818.89	797.34	21.55	1.02703	0.00129	1.35	0.0522
Trifluoroethanol	Kerosene	1380.57	797.34	583.23	1.73147	0.00129	8.32	0.0069
Ethylene glycol	Kerosene	1100.01	797.34	302.67	1.37960	0.00129	18.30	0.0099
Dimethylformamide	Hexane	947.15	665.37	281.78	1.42349	0.00032	4.69	0.0061
Acetonitrile	Hexane	778.26	665.37	112.89	1.16966	0.00032	5.36	0.0105
Ethylene glycol	Hexane	1100.01	665.37	434.64	1.65323	0.00032	25.55	0.0048
Water	Hexane	1000.00	665.37	334.63	1.50292	0.00032	17.44	0.0056
Water	Xylene	1000.00	857.95	142.05	1.16557	0.00060	12.48	0.0123

r = radius of the liquid drop; u = terminal velocity of the drop; ρ = relative density of the liquid drop; σ = relative density of the liquid in the column; η = viscosity of the liquid in the column; γ = interfacial tension between the liquid drop and the liquid in the column.

where K is a constant and π_2 is the reciprocal of the Reynolds number Re and $\pi_3=\rho/\sigma$, the ratio of the densities. Subsituting the values of π_1, π_2, and π_3 in eq. (1), the expression for drag force is

$$F = KD^2u^2\sigma(\eta/\sigma uD)^{1/2}(\rho/\sigma)^{1/2},$$

or

$$F = KD^{3/2}u^{3/2}\eta^{1/2}\rho^{1/2}. \quad (2)$$

Since at the terminal velocity

$$F = 4\pi r^3(\rho - \sigma)g/3, \quad (3)$$

where r is the radius of the drop and is equal to $D/2$, we get the constant K as

$$K = \sqrt{2}\pi g(r/u)^{3/2}(\rho - \sigma)/3\eta^{1/2}\rho^{1/2}, \quad (4)$$

or the terminal velocity u as

$$u = (2\pi^2g^2r^3(\rho - \sigma)^2/9\eta K^2\rho)^{1/3}. \quad (5)$$

Rewriting eq. (4), we get

$$[(r/u)^{3/2}(\rho - \sigma)]/[\eta^{1/2}\rho^{1/2}] = 3K/\sqrt{2}\pi g, \quad (6)$$

i.e. $$\sqrt{2}\pi g\, S/3K = 1, \quad (7)$$

where

$$S = [(r/u)^{3/2}(\rho - \sigma)]/[\eta^{1/2}\rho^{1/2}]. \quad (8)$$

Equation (8) is of vital importance, in the sense that it can be used to determine the density (ρ) of liquids available in microquantities when the conventional methods fail. S has been found to be approximately a constant for all systems (see Table 3), the mean value of which is 0.248319 m^{-1} s^2.

Equation (8) may be written as a quadratic equation which gives the value for ρ as

$$\rho = \{2\sigma + \lambda + [\lambda(\lambda + 4\sigma)]^{1/2}\}/2, \quad (9)$$

where

$$\lambda = S^2\eta/(r/u)^3. \quad (10)$$

Equation (9) may be employed to determine the density of liquids available in microquantities.

A graduated cylinder (diameter 5 cm) filled with the experimental liquid was used. Liquid drops of known volume were gently injected into the liquid column using a graduated Hamilton precision microsyringe. The drops and the liquid were immiscible. The terminal

Table 2. Experimental data for liquid drop pairs

Liquid drop pair	Volume V of the drop (l)	Radius r of the drop (10^{-4} m)	Distance S travelled (10^{-2} m)	Time t taken (s)	Observed terminal velocity u (10^{-2} m s^{-1})	r/u (s)	Reynolds number	Eotvos number ($\times 10^{-2}$)
1-2-Dichlorobenzene in water	0.5	4.9237	80	16.4	4.8780	0.0101	48.04	3.280
Morton no. = 4.7403×10^{-12}	1.0	6.2035	50	8.1	6.1728	0.0100	76.59	5.207
	2.0	7.8159	50	6.4	7.8125	0.0100	122.12	8.265
1-2-Dichloroethane in water	0.5	4.9237	40	9.1	4.3956	0.0112	43.29	7.263
Morton no. = 7.4139×10^{-11}	1.0	6.2035	40	7.2	5.5556	0.0112	68.93	11.529
	2.0	7.8159	40	5.7	7.0175	0.0111	109.70	18.301
Water in kerosene	0.5	4.9237	40	10.2	3.9216	0.0126	23.87	4.399
Morton no. = 1.0307×10^{-10}	1.0	6.2035	40	8.1	4.9383	0.0126	37.87	6.983
	2.0	7.8159	40	6.4	6.2500	0.0125	60.39	11.084
Ethylene glycol in kerosene	0.5	4.9237	40	8.0	5.0000	0.0098	30.43	15.779
Morton no. = 2.1089×10^{-9}	1.0	6.2035	40	6.4	6.2500	0.0099	47.93	24.953
	2.0	7.8159	40	5.1	7.8431	0.0100	75.78	39.610
Dimethylformamide in	0.2	3.6278	40	6.8	5.8824	0.0062	88.74	30.988
hexane	0.3	4.1528	40	5.8	6.8966	0.0060	119.10	40.606
Morton no. = 6.3348×10^{-10}	0.4	4.5708	40	5.3	7.5472	0.0061	143.46	49.191
Acetonitrile in hexane	0.2	3.6278	40	11.5	3.4783	0.0104	52.47	10.857
Morton no. = 1.6974×10^{-10}	0.3	4.1528	40	10.1	3.9604	0.0105	68.39	14.227
	0.4	4.5708	40	9.2	4.3478	0.0105	82.64	17.235
Water in xylene	0.5	4.9237	40	10.0	4.0000	0.0123	56.32	10.818
Morton no. = 1.2613×10^{-10}	1.0	6.2035	40	7.9	5.0633	0.0123	89.83	17.172
	2.0	7.8159	40	6.3	6.3492	0.0123	141.92	27.259

Table 3. The value of $(r/u)^{3/2}/\eta^{1/2}$ and $\rho^{1/2}/(\rho - \sigma)$ and the constant K obtained by using eq. (6)

Liquid drop	Column liquid	$(r/u)^{3/2}/\eta^{1/2}$	$\rho^{1/2}/(\rho - \sigma)$	$S^{\dagger}$	K^{*}
Eugenol	Water	0.11392	0.44985	0.253239	3.6754
Chlorobenzene	Water	0.08611	0.33816	0.254644	3.6958
Benzoyl benzoate	Water	0.07510	0.29531	0.254302	3.6908
1-2-Dichloroethane	Water	0.03698	0.14514	0.254806	3.6981
1-2-Dichlorobenzene	Water	0.03162	0.12327	0.256533	3.7232
Trichloroethylene	Water	0.02137	0.08558	0.249676	3.6237
Chloroform	Water	0.01972	0.08051	0.244972	3.5554
Bromobenzene	Water	0.01892	0.07848	0.241058	3.4986
Carbon tetrachloride	Water	0.01657	0.06857	0.241681	3.5076
Nitrobenzene	Water	0.04419	0.17818	0.247730	3.5997
Carbon disulphide	Water	0.03500	0.14056	0.249707	3.6139
Dichloromethane	Water	0.02882	0.11534	0.248279	3.6265
Water	Kerosene	0.03891	0.15604	0.250319	3.6192
Alcohol	Kerosene	0.33206	1.32790	0.249674	3.6292
Trifluoroethanol	Kerosene	0.01582	0.06371	0.248203	3.6039
Ethylene glycol	Kerosene	0.02743	0.10958	0.249257	3.6324
Dimethylformamide	Hexane	0.02663	0.10922	0.241214	3.5391
Acetonitrile	Hexane	0.06015	0.24712	0.242377	3.5324
Ethylene glycol	Hexane	0.01859	0.07631	0.242101	3.5358
Water	Hexane	0.02343	0.09450	0.244846	3.5978
Water	Xylene	0.05569	0.22262	0.250073	3.6307
			Mean	0.248319	3.6109

*The constant of eq. (6)
†S of eq. (8)

velocity u was determined by noting the time of transit t required by the liquid drop of radius r to cover a distance S between two graduations on the column. The drops studied here while falling are ellipsoidal in shape. If V is the volume of the drop of equivalent radius[7] r, then $r = (3V/4\pi)^{1/3}$.

The densities of the liquids were determined accurately by the specific gravity bottle method (Table 1). Viscosity values were determined by using the Ostwald viscometer (Table 1) and interfacial tension values have been determined by the method of drops (Table 1).

Twenty-one liquid systems were studied (Table 1). Only seven liquid drop pairs having three data points each are presented to show that r/u is approximately constant (Table 2). The other fourteen liquid drop pairs for which experimental results have been obtained satisfy the same. The right-hand side of eq. (6) is a constant. A graph is drawn with $(r/u)^{3/2}/\eta^{1/2}$ along the X-axis and $\rho^{1/2}/(\rho - \sigma)$ on the Y-axis (Figure 1). The slope of the line in the graph gives the value of $\sqrt{2\pi g/3K}$.

Substituting the experimental value (0.248319) obtained by eq. (8) (Table 3) on the left-hand side of eq. (6), the value of K can be determined. The value of K thus determined (Table 3) is 3.6109, and that determined from the graph (Figure 1) is 3.6079, the mean being 3.6094. Therefore, for the motion of a liquid drop falling freely and vertically, and not exhibiting oscillation, through an immiscible liquid medium of fairly high Reynolds, Eotvos and Morton numbers, eq.(2) is the fresh drag force expression, in which the experimental mean value of the constant K is 3.6094.

Equation (5) with $K = 3.6094$ may be used to predict the terminal velocity of the drop falling without oscillation. Using eq. (8) (with mean experimental value 0.248319, Table 3), if the terminal velocity and radius are known, one can determine the density of a micro-quantity liquid drop for which the density cannot be determined by the capillary tube method (where weighing is a problem) or by any other conventional methods.

1. Acheson, D. J., *Elementary Fluid Dynamics*, Clarendon Press, Oxford, 1990, pp. 226.
2. Batchelor, G. K., *An Introduction to Fluid Dynamics*, American Book Company, New York, 1967, pp. 75–101.
3. Kay, J. M. and Nedderman, R. M., *Fluid Mechanics and Transfer Processes*, University Press, Cambridge, 1985, pp. 144.
4. Massey, B. S., *Mechanics of Fluids*, Van Nostrand Reinhold, London, 1989, pp. 261.
5. Smith, C. J., *Intermediate Physics*, Arnold, London, 1959, pp. 137.
6. Allen, H. S., *Phil. Mag.*, 1900, **5**, 323–333, 519–533.
7. Clift, R., Grace, J. R. and Weber, M. E., *Bubbles, Drops and Particles*, Academic Press, New York, 1978.
8. Davis, R. E. and Acrivos, A., *Chem. Engg. Sci.*, 1966, **21**, 681–685.
9. Edge, R. M. and Grant, C. D., *Chem Engg. Sci.*, 1972, **27**, 1709–1721.
10. Elzinga, Jr, E. R. and Banchero, J. T., *AICHE J.*, 1961, **7**, 394–399.
11. Garner, F. H. and Skelland, A. H. P., *Chem. Engg. Sci.*, 1955, **4**, 149–158.
12. Gibbons, J. H., Houghton, G. and Coull, J., *AICHE J.*, 1962, **8**, 274–276.
13. Grace, J. R., Wairegi, T. and Nguyen, T. H., *Trans. Inst. Chem. Eng.*, 1976, **54**, 167–173.
14. Griffith, R. M., *Chem. Engg. Sci.*, 1962, **17**, 1057–1070.
15. Harmathy, T. Z., *AICHE J.*, 1960, **6**, 281–288.
16. Hendrix, C. D., Dave, S. B. and Johnson, H. F., *AICHE J.*, 1967, **13**, 1072–1077.
17. Hu, S. and Kinter, R. C., *AICHE J.*, 1955, **1**, 42–48.
18. Johnson, A. I. and Braida, L., *Can. J. Chem. Engg.*, 1957, **35**, 165–172.
19. Kintner, R. C., *Advances in Chemical Engineering*, Academic Press, New York, 1963, pp. 51–94.
20. Klee, A. J. and Treybal, R. E., *AICHE J.*, 1956, **2**, 444–447.
21. Krishna, P. M., Venkateswarlu, D. and Narasimhamurty, G. S. R., *Chem. Engg. Data*, 1959, **4**, 336–343.
22. Licht, W. and Narasimhamurty, G. S. R., *AICHE J.*, 1955, **1**, 366–373.
23. Mendelson, H.D., *AICHE J.*, 1967, **13**, 250–253.
24. Raymond, D. R. and Zieminskei, S. A., *AICHE J.*, 1971, **17**, 57–65.
25. Skelland, A. H. P. and Caenepeel, C. L., *AICHE J.*, 1972, **18**, 1154–1163.
26. Thorsen, G., Stordalen, R. M. and Terjesen, S. G., *Chem. Engg. Sci.*, 1968, **23**, 413–426.
27. Wellek, R. M., Agarwal, A. K. and Skelland, A. H. P., *AICHE J.*, 1966, **12**, 854–862.
28. Winnikow, S. and Chao, B. T., *Phys. Fluids*, 1966, **9**, 50–61.
29. Yeheskel, J. and Kehat, E., *Chem. Engg. Sci.*, 1971, **26**, 1223–1233.
30. Zabel, T., Hanson, C. and Ingham, J., *Trans. Inst. Chem. Eng.*, 1973, **51**, 162–164.

Received 9 December 1994; revised accepted 19 July 1995

Antibody response to a potentially protective antigen in human filariasis

Madhusmita Bal and Manoj K. Das

Parasite Immunology Division, Regional Medical Research Centre (ICMR), Nandan Kanan Road, Bhubaneswar 751 016, India

Antibody isotype (IgG and IgM) levels to a high molecular weight antigen (1.56×10^5 kDa) isolated from water-insoluble components of *Setaria digitata* were determined to be highly elevated in microfilariae-negative individuals of a *Wuchereria bancrofti* endemic area of Orissa, India. The highest IgM response was found in endemic normals. The prevalence of antibodies was minimum in microfilaraemic group and maximum in endemic normals of the population.

LYMPHATIC filariasis is one of the most prevalent parasitic diseases, estimated to have infected 78 million people globally. In India alone about 45 million people are infected primarily by *Wuchereria bancrofti*[1]. It is a chronic debilitating disease of humans initiated by infective bites of mosquitoes *Culex quinquefasciatus*.

Much attention has been focused recently on the concept of protective immunity in filariasis[2,3]. Various models of animal filariasis indicate that protective responses can be induced by exposure to live or irradiated parasites[4,5]. However, very little is known about protective immunity in human filariasis and there is no direct evidence that such immunity exists in humans. On the other hand, epidemiological studies indicate that age-related microfilarial rates increase and then decrease in older-age classes of endemic population[6]. A proportion of adult residents remain amicrofilaraemic despite intensive exposure to infective larvae (L_3) of filarial parasites. Such disease-free adult residents (endemic normals) are believed to have developed protective immunity to filarial infection. It is presumed that host protective immunity might be effective against the acquisition of (new) infection. Identification of antigens associated with the resistance to infection under natural conditions will be of great importance in deciphering the complex mechanism of filarial protective immunity. In the present report we describe a high molecular weight antigen of *Setaria digitata* which stimulates increased antibody responses only in amicrofilaraemic individuals. Moreover, the IgM response to this antigen is considerably elevated, especially in endemic normals compared to infected individuals.

The study population is from *Wuchereria bancrofti* infected village, Baniatangi (Khurda district, Orissa), about 45 km away from Bhubaneswar, Orissa. Individuals were classified into different clinical groups, namely asymptomatic microfilaraemics (AS), filariasis with chronic pathology elephantiasis and/or hydrocele (CP) and endemic normals (EN) as described earlier[7]. In a random survey of 318 people (age 3–75 years) in the village, AS and CP patients constituted 15.09 and 13.20% of the population, respectively. No attempt was made to include individuals with acute filarial symptoms (adenolymphangitis). Disease-free microfilariae (MF) negative adults (≥ 18 years, 21.83%) were regarded as endemic normals of the region.

Setaria digitata adults were collected from cattle in the local slaughterhouse. Aqueous-insoluble materials were obtained as a pellet by centrifuging the extracts of worms following homogenization and sonication in saline. The pellet was solubilized in the detergent NP-40 (0.5% NP-40 in 0.01 M Tris–HCl, pH 8.0) by keeping for 2 h at 25°C with occasional shaking. The extract was centrifuged again and ammonium sulphate (50%) was added to the supernatant, which was kept at 4°C overnight. The precipitate collected after centrifugation was dissolved in water and was dialysed against phosphate-buffered saline PBS (0.01 M phosphate, 0.15 M NaCl, pH 7.2). The dialysed material upon Sephadex G-200 column chromatography resolved into a major peak (designated as $DSSd_1$) with a molecular mass of 1.56×10^5 kDa, eluted after the void volume, followed by a minor peak.

Pooled sera were made by adding equal volumes of serum from individuals ($n = 40$ in each category) belonging to different categories of bancroftian filariasis. A control serum pool was also prepared from 20 residents of a nonfilarial region (Koraput district) of Orissa and was used as a nonendemic (NEN) sample. Levels of serum antibody to $DSSd_1$ were measured by ELISA using 96-well polysterene plates. The wells were coated by incubation at 37°C (5 h) and then 4°C (overnight) with antigen (3 μg/ml) in bicarbonate buffer (pH 9.6). Next, the wells were blocked with 0.4% bovine serum albumin (BSA) at 37°C (1 h), then incubated (3 h) with diluted serum samples in Tween-20 (0.1% in phosphate buffer saline (PBS)). Final incubations were performed for 1 h at 37°C and 4°C (overnight) with peroxidase-conjugated rabbit anti-human IgG or IgM (Dakopatts) diluted to 1:1000 in 0.1% Tween-20, followed by colour development using orthophenylenediamine substrate. Absorbance was read at 492 nm using an ELISA plate reader (Bio-Rad).

Figure 1 depicts the antibody levels to $DSSd_1$ in different dilutions of pooled sera. It can be seen clearly that AS serum among the filarial groups exhibited the lowest antibody response. Both EN and CP sera had high antibody levels. IgM levels were highest in EN serum, while CP serum had highest levels of IgG. Antibodies in control NEN serum were found to be negligible at 1:400 dilution. IgM and IgG levels of individual sera were determined (Figure 2). The mean antibody values and the extent of sero-positivity (prevalence) are shown in Table 1 and Figure 3.

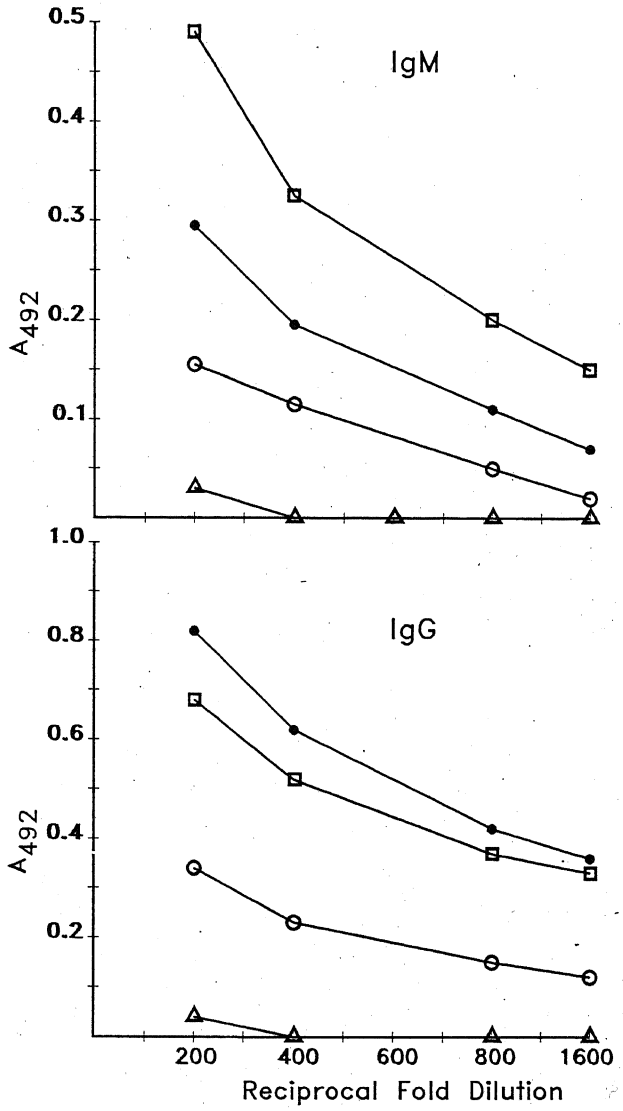


Figure 1. IgM and IgG antibody levels to $DSSd_1$ in pooled filarial sera: chronic filariasis (•); asymptomatic microfilaraemic carriers (o); endemic normals (□) and nonendemic normals (Δ).

The prevalence of $DSSd_1$-specific antibodies in filarial groups was evaluated (Figure 3). IgG and IgM prevalence in AS group is low; only 27.5 and 10%, respectively. The prevalence of these two isotypes in EN is 75 and 85% and in CP 100 and 32%, respectively. All individuals in CP are IgG-positive but only 32% of CP patients have IgM antibodies. EN as a group have the highest of prevalence for both isotypes.

An interesting approach to study the protective immunity is to analyse the immune response of endemic normal ('immune') individuals *vis-à-vis* infected persons. Endemic normals can be assumed as putatively immune since they fail to acquire a detectable parasite burden (MF load and/or filarial symptoms). Such an approach has led to the identification of a 43 kDa larval antigen that is recognized only by endemic normals of a

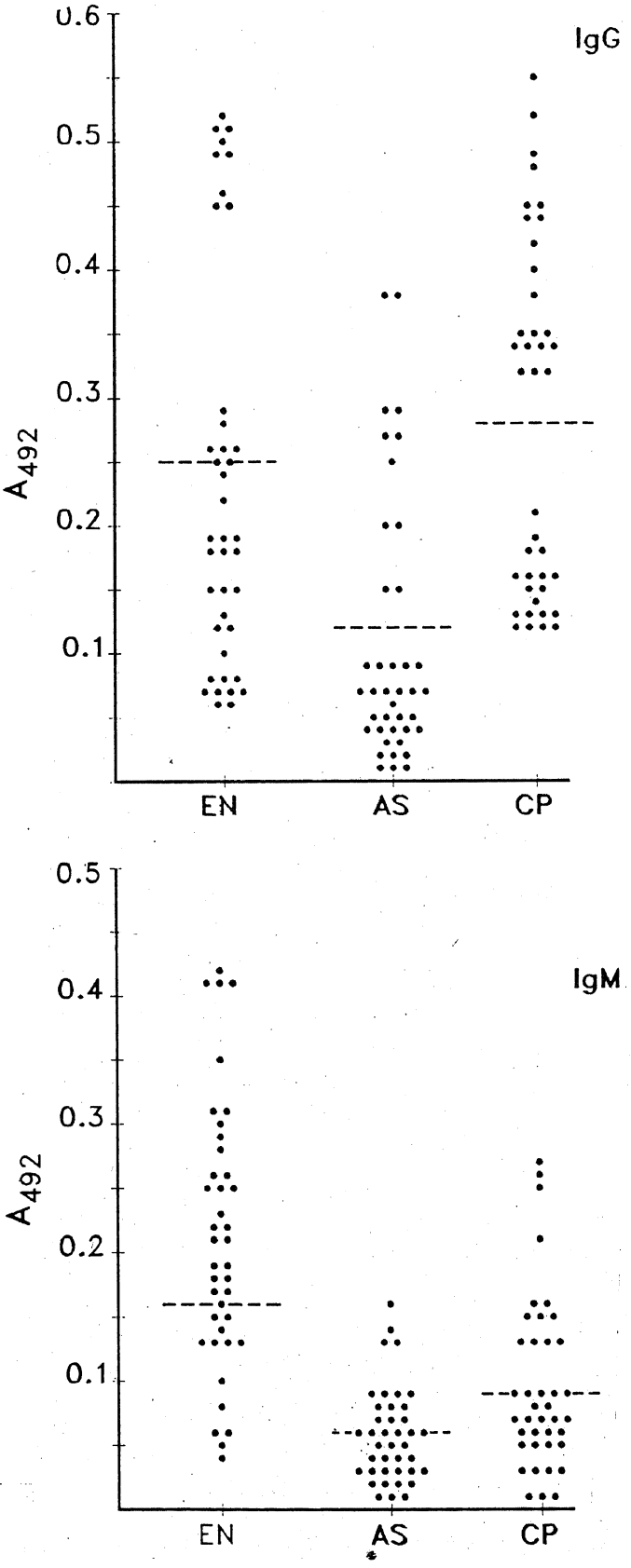


Figure 2. IgG and IgM levels to $DSSd_1$ antigens in filarial sera. The horizontal bar indicates the mean values.

Table 1. ELISA analysis of $DSSD_1$-specific antibody response. Each sample ($n = 40$ in each filarial group) was tested in triplicate (at 1:400 dilution)

Group	IgM		IgG	
	OD 492 (range)	Percentage positive (+ n)	OD 492 (range)	Percentage positive (+ n)
Endemic normals	0.16 ± 0.08* (0.03–0.41)	85 (34)	0.25 ± 0.16* (0.07–0.51)	75(30)
Asymptomatic microfilaraemic	0.06 ± 0.03 (0.01–0.14)	10(4)	0.12 ± 0.10 (0.01–0.38)	27.5(11)
Chronic filariasis	0.09 ± 0.06** (0.01–0.27)	32.5(13)	0.28 ± 0.14* (0.12–0.58)	100(400)
Nonendemic normals	0.03 ± 0.03 (0–0.07)	0	0.04 ± 0.02 (0.02–0.09)	0

*$P < 0.01$ compared to asymptomatic microfilaraemics.

**$P < 0.01$ compared to endemic normal.

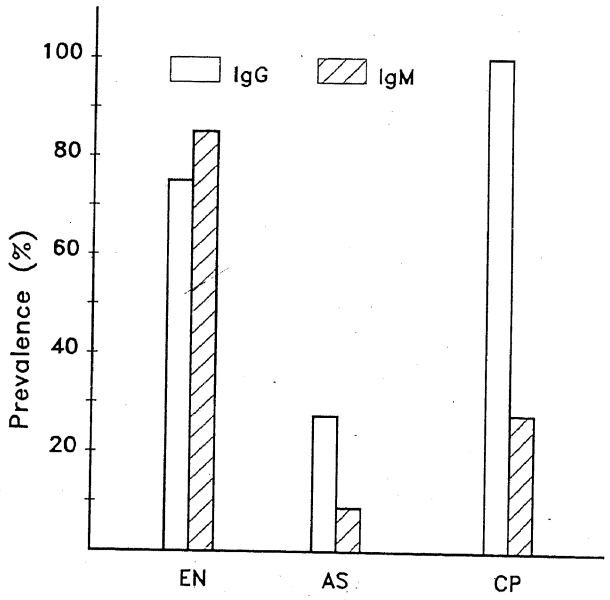


Figure 3. The prevalence of antibodies to $DSSd_1$ in filarial endemic region. Isotypic seropositivity was calculated at 1:400 dilution with respect to nonendemic normals (> mean + 3 SD; n=20). The number of sera in each filarial group was 40.

filarial endemic region[8]. Similarly, in schistosomiasis, infection-free individuals have higher levels of antibodies to paramyosin than the infected persons[9]. Interestingly, paramyosin has also been shown as a protective antigen in *B. malayi* infected animals[10,11]. Several other filarial antigens with protective potential have been identified and characterized[12–14], but studies on human immune responses are limited.

Two notable findings of the present report are: (1) the depressed antibody responses to $DSSd_1$ in microfilaraemic carriers AS group and (2) the elevated antibody response in EN, even higher than in the CP group. This is also reflected in antibody prevalence, which was found to be low in AS, rising to higher values in amicrofilaraemics EN and CP groups. These observations become more marked in comparison to NEN individuals (Figure 1) and suggest that the antibody response that was measured here is specific to filarial infection. The antigen $DSSd_1$ was isolated from the aqueous-insoluble materials of filarial parasite which are normally discarded. In this sense this antigen might constitute an hitherto undetected novel target.

The presence of increased antibody levels in amicrofilaraemic individuals gives rise to the possibility that $DSSd_1$ antigen could be a target of antimicrofilarial immunity in the course of natural infection.

1. World Health Organization, Technical Report series 821, Fifth Report of the WHO Expert Committee on Filariasis, 1992.
2. Nutman, T. B., *Exp. Parasitol.*, 1989, **68**, 248–252.
3. Selkirk, M. E., Maizels, R. M. and Yazdanbakksh, M., *Immunobiology*, 1992, **184**, 263–281.
4. Hayashi, Y., Nagasaki, S., Hammerberg, B. and Tanaka, H., *Am. J. Trop. Med. Hyg.*, 1989, **41**, 650–656.
5. Abraham, D., Grieve, R. B., Holy J. M. and Christensen, B. M., *Am. J. Trop. Med. Hyg.*, 1989, **40**, 598–604.
6. Bundy, D. A. F., Grenfell, B. T. and Rajgopalan, P. K., *Immunoparasitol. Today*, 1991, **10**, A71–A75.
7. Das, M. K., Beuria, M. K. and Dash, A. P., *Int. Arch. Allergy Immunol.*, 1992, **99**, 118–122.
8. Freedman, D. D., Nutman, T. B. and Ottesen, E. A., *J. Clin. Invest.*, 1989, **83**, 14–22.
9. Correa-Oliveria, R. C., Pearce, E. J., Liveria, G. C., Golgher, D. B., Katz, N., Bahia, I. G., Carvalho, O. S., Gazzinelli, G. and Sher, A., *Trans. R. Soc. Trop. Med. Hyg.*, 1989, **83**, 798–804.
10. Nanduri, J. and Kazura, J. W., *J. Immunol.*, 1989, **143**, 3359–3363.
11. Li, B. W., Chandrashekhar, R., Alvarez, R. M., Liftis, F. and Weil, G. J., *Mol. Biochem. Parasitol.*, 1991, **49**, 315–324.
12. Kazura, J. W., Cicirello, H. and Forsyth, K., *J. Clin. Invest.*, 1986, **77**, 1988–1992.
13. Rothstein, N. M., Higashi, G., Yates, J. and Rajan, T. V., *Mol. Biochem. Parasitol.*, 1989, **33**, 229–236.
14. Kazura, J. W., Hazlet, F. E., Pearlman, E., Day, K., El Zeiny, A., Nilsen, T. W. and Alper, M. P., *J. Infect. Dis.*, 1992, **166**, 1453–1457.

ACKNOWLEDGEMENTS. We thank Dr K. Satyanarayana for interest. The residents of Baniatangi village are acknowledged for their cooperation.

Received 12 May 1995; revised accepted 3 August 1995

Visual demonstration of *in vivo* acid phosphatase activity of VA mycorrhizal fungi

J. C. Tarafdar
Central Arid Zone Research Institute, Jodhpur 342 003, India

For nondestructive and visual demonstration of the release of acid phosphatase by vesicular–arbuscular (VA) mycorrhizal fungi, filter papers treated with a mixture of 1-naphthyl phosphate as substrate and the diazonium salt Fast Red TR as an indicator were used. Wheat (*Triticum aestivum* L.) was grown for 6 weeks in sterilized soils in specially designed pots with five compartments with or without hyphal and root barriers. Plants were inoculated with *Glomus mosseae* (Nicol & Gerd) Gerd & Trappe and the treated filter paper was placed at the outer surface of the hyphal compartment. Acid phosphatase activity was visualized as a red-coloured 'hyphal print' on the filter paper.

PHOSPHATASE activity in the rhizosphere may originate from the plant roots[1], from fungi such as *Aspergillus*[2], ectomycorrhizas[1] or from bacteria[3]. Vesicular–arbuscular (VA) mycorrhizas are widespread in soils, and their beneficial effect on plant growth in P-deficient soils is attributed to uptake and translocation of P by external hyphae[4]. VA mycorrhizal storage of polyphosphates in the hyphae has also been reviewed[5]. Our recent study indicated[6] the production of phosphatase by external hyphae. Here a nondestructive method for the visualization of acid phosphatase activity by VA mycorrhizal fungal hyphae is described.

A special Plexiglass pot (Figure 1) comprising of five compartments, a central one for root growth, two adjacent compartments separated from the central compartment by a nylon net of 30 μm mesh size, for growth of VA mycorrhizal hyphae (hyphal compartments), and two other compartments, separated by a 0.45 μm membrane, through which neither roots nor hyphae could penetrate, has been described by Tarafdar and Marschner[6]. It was modified by placing a Plexiglass barrier (2 cm × 2 cm) between hyphae and outer compartment on one side of the pot to prevent hyphae to pass into that side. Wheat plants (cv. Star) were grown in sterilized soils of Bavendorf under controlled conditions. The soil used was a cambisol of pH 6.9 and sieved (2 mm) prior to sterilization (48 h each at 120°C, on two alternate days). Water was supplied directly into the root compartment using PVC tubes with string wicks inserted in bottles containing sterilized distilled water.

The mycorrhizal fungus used was *Glomus mosseae*. Each pot received about 1500 surface-sterilized (with 0.2% chloramin T and 0.02% streptomycin sulphate) infective propagules of 90–250 μm in size. The inoculum was mixed uniformly with the whole soil, under aseptic conditions, and added to the central (root) compartment of the pots. Wheat seeds were surface-sterilized and planted, four plants in each pot, into the root compartment. The top openings of all compartments were covered with 2 cm layer of sterilized quartz sand. All operations were performed in a laminar flow chamber. The pots were then transferred to a growth chamber (25–18°C and 15–9 h day–night regime, 305 μmol m^{-2} s^{-1} photon flux density).

For demonstration of phosphatase activity the method used for determination of acid phosphatase in blood serum has been modified by soaking filter papers with the reagents. In this method enzymatic hydrolysis of 1-naphthyl phosphate at pH 5.6 liberates 1-naphthol which reacts with the diazonium salt Fast Red TR (diazotized 2-amino-5 chlorotoluene 1,5-naphthalene disulphonate), forming a red dye. The suitability of 1-naphthyl phosphate as a substrate for phosphatase activity from plant tissues was tested[1] with acid phosphatase isolated from wheat germ.

For the assay the following solutions were prepared – citrate buffer: trisodium citrate solution (50 mM), adjusted to pH 5.6 with HCl (1 M); substrate solution: 37.5 mM 1-naphthyl phosphate sodium salt in citrate buffer; Fast Red TR solution: 2.7 mM Fast Red TR solution in citrate buffer. The substrate solution and the Fast Red TR solution were mixed in a ratio of 1:10 (v/v) and poured into a Petri dish. After soaking in filter papers in the assay mixture for some seconds, they were placed on Plexiglass and the excess solution was removed with paper tissues.

The impregnation and air drying of the filter papers should be carried out in a fumehood to prevent the inhalation of toxic volatiles and also to accelerate the drying process. Thereafter, a 5 mm margin of the filter papers is cut off, as it contains higher concentrations of reagents than the remaining filter paper. The filter papers should be prepared no longer than 2 days before application and stored by wrapping in an aluminium foil in a desiccator. Ashless filter paper (Blue ribbon, No. 589, Schleicher & Schüll, Dassel/Germany) was used.

For the visualization of acid phosphatase activity the pots were opened and the outer compartments separated, and an impregnated filter paper was placed on the surface of the mesh outside the hyphal compartment on both the sides (Figure 1). All the tests were carried out in the growth chamber at 25°C. The filter papers were removed after different times of application and dried.

Acid phosphatase activity by VAM hyphae is visualized as dark red (black on the photographs) 'hyphal print' on the test papers (Figure 2 *b–c*). All parts of the hyphae produced a more or less intensive hyphal print. When hyphae were not allowed to pass by introducing a block, no hyphal print in that area could be

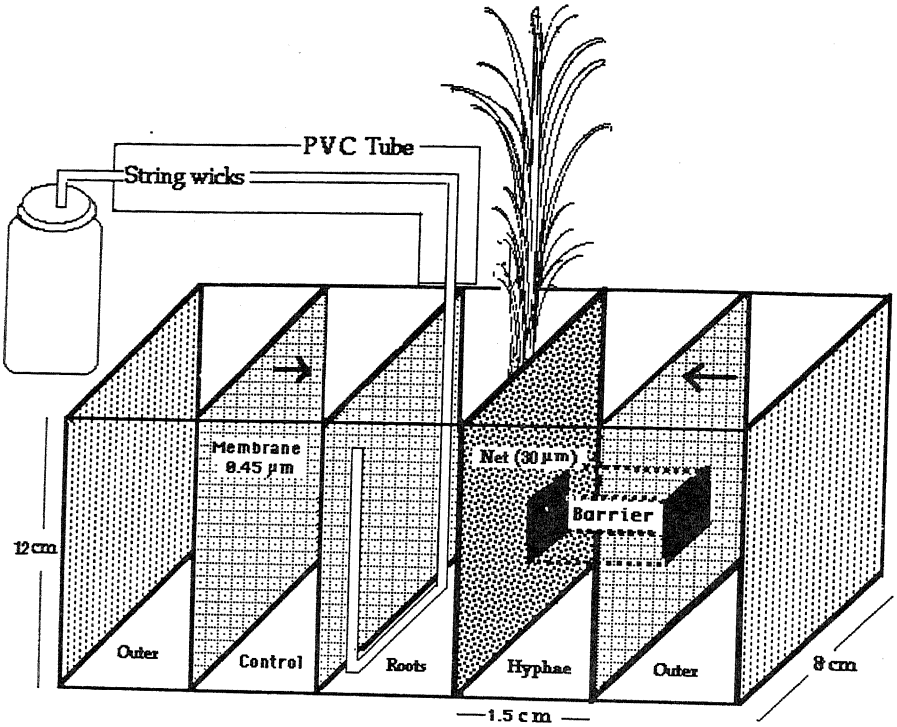


Figure 1. Details of the pots used in the experiment (modified after ref. 6). Arrow denotes the position where the filter paper was put.

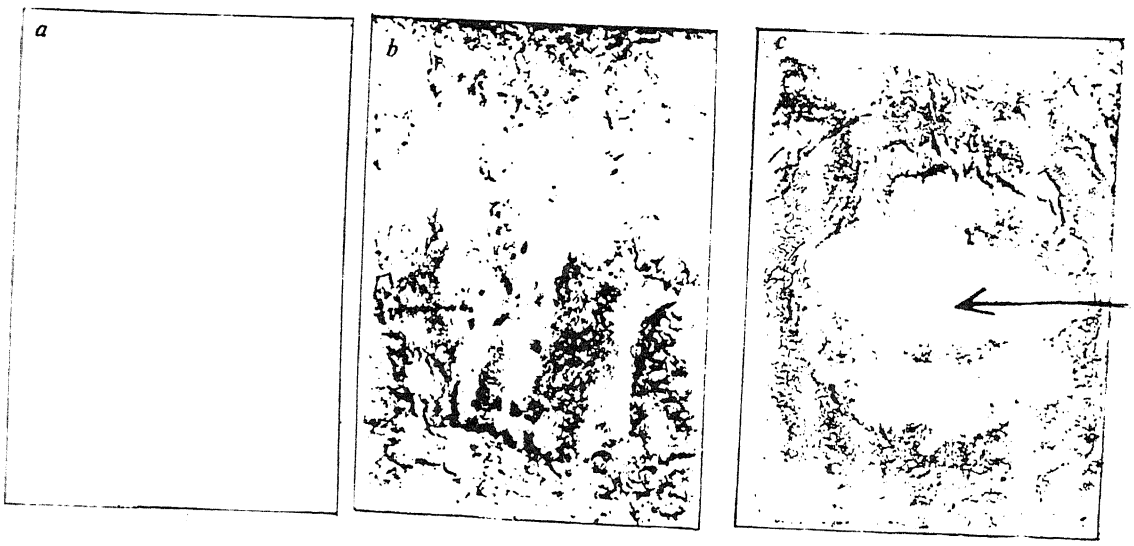


Figure 2. Fingerprints of VA mycorrhizal hyphae: *a*, control; *b*, +VAM; *c*, +VAM with barrier. Arrow denotes the position where the barrier was put on the hyphal path.

seen (Figure 2 *c*). Similarly, no impression was noticed in control (uninoculated pots, Figure 2 *c*). Therefore, these hyphal prints mainly originated from the activity of hyphal surface phosphatase. The dark colour indicates a high phosphatase activity of these hyphal zones. As expected, the intensity of the colour of the hyphal prints depends on the duration of application. Application for only 10 min resulted in a faint, not well-differentiated colouration, whereas it got intensified gradually, with maximum intensity showing when exposed for 60 min.

Acid phosphatase activity in plant roots and soils is usually measured with *p*-nitrophenyl phosphate (pNPP) as the substrate. Since the end product, *p*-nitrophenol is soluble, spraying the filter paper with an alkaline solution at the end of the application period may result in diffusion of the dye and the formation of blurred hyphal prints. In contrast, 1-naphthyl phosphate results in an insoluble end product when no detergent is used. Thus, sharp hyphal prints appear on the filter paper. In addition, the method with 1-naphthyl phosphate has the advantage that the reaction between 1-naphthol and Fast Red TR takes place at low pH. Since acid phosphatases are characterized by maximal activities at low pH values[7], the method with 1-naphthyl phosphate permits a direct and continuous measurement of acid phosphatase activity at its optimal pH.

While 1-naphthyl phosphate is frequently used for the detection of acid phosphatase in medical media[7], only a few reports exist on the use of this substrate for the determination of acid phosphatase from soils[8,9] and from plant roots[1,9]. *In vitro* assay with acid phosphatase from wheat germ also shows that 1-naphthyl phosphate is a very suitable substrate. The present experiment confirmed the production of phosphatase by VA mycorrhizal hyphae. The method may be used for the comparison of phosphatase activity of different VA myocorrhizal hyphae as well as plant species or genotypes in relation to their capacity to utilize organic phosphorus depending upon plant age, phosphorus nutritional status, temperature and other environmental factors.

1. Dinkelaker, B. and Marschner, H., *Plant Soil*, 1992, **144**, 199–205.
2. Tarafdar, J. C., Rao, A. V. and Bala, K., *Folia Microbiol.*, 1988, **33**, 453–457.
3. Tarafdar, J. C. and Claassen, N., *Biol. Fertil. Soils*, 1988, **5**, 308–312.
4. Li, X-L., George, E. and Marschner, H., *Plant Soil*, 1991, **136**, 41–48.
5. Gianinazzi-Pearson, V. and Gianinazzi, S., in *Physiological and Genetical Aspect of Mycorrhizae* (eds Gianinazzi-Pearson, V. and Gianinazzi, S.), INRA, Paris, 1986, pp. 101–109.
6. Tarafdar, J. C. and Marschner, H., *Soil Biol. Biochem.*, 1994, **26**, 387–395.
7. Moss, D. W., in *Methods of Enzymatic Analysis* (eds Bergmeyer, J. and Graßl, M.), Verlag Chemie, Weinheim, 1984, vol. IV, pp. 92–106.
8. Ramirez-Martinez, J. R. and McLaren, A. D., *Enzymology*, 1966, **30**, 243–253.
9. Tarafdar, J. C., *J. Indian Soc. Soil Sci.*, 1989, **37**, 393–395.

ACKNOWLEDGEMENT. This research was financially supported by a grant from the Alexander von Humboldt Foundation (AvH) and was done at Institut für Pflanzenernährung, Universität Hohenheim, Früwirthstraße 20, 70599 Stuttgart, Germany.

Received 28 December 1994; revised accepted 5 August 1995

In vitro flowering and pod formation in cauliflower (*Brassica oleracea* var. *botrytis*)

Vandana A. Kumar, Atul Kumar and Jitendra Kumar

College of Forestry and Hill Agriculture, G.B. Pant University of Agriculture and Technology, Hill Campus, Tehri Garhwal 249 199, India

Flowering and pod formation *in vitro* were induced in cauliflower through culture of curd explants in Murashige and Skoog medium supplemented with 3.0 mg l^{-1} indole acetic acid (IAA) and 0.5 mg l^{-1} kinetin after incubation under 16 h photoperiod at 28 ± 2°C temperature during the day and 24 ± 2°C during the night at a light intensity of 3000 lx. Pods were ready to harvest within 90 days of culture initiation and developed small, globular, smooth and brown-coloured seeds. This phenomenon, though gives small amount of seeds, may be helpful when pollination and fertilization *in vivo* are difficult. This way, the long period for flowering and pod formation under field conditions is cut short and the problem of isolation distance in cross-pollinated crops (like cauliflower) is also overcome.

BREEDING programmes in cauliflower rely mainly on raising seed crops and, under field conditions, proper isolation distance between varieties is necessary for hybridization. Besides, curding and flowering are also very sensitive to changes in temperature[1]. Initiation of flowering and complete pod formation *in vitro* may become a valuable research tool for plant breeders ensuring seed purity. This can also facilitate hybridization *in vitro* without any requirement for isolation distance. A protocol for *in vitro* flowering and pod formation in cauliflower is described.

Florets of 1 cm diameter were excised from a mature curd of cauliflower (snowball-type), surface-sterilized with 70% (v/v) ethanol for 90 s, followed by commercial bleach (calcium hypochlorite) solution (5%, w/v) containing 0.1% Tween-20 for 15 min. After three rinses with sterile distilled water, small (approx. 3 mm cubic) pieces of curd surface were dissected and implanted in the

Table 1. *In vitro* plantlet regeneration in cauliflower

Treatment (mg l^{-1})	Rooting after 30 days		Plantlet formation after 30 days
	No. of roots	Length of roots (cm)	No. of developed leaves
IAA 0.5 + kinetin 0.5 (A)	3.4 ± 0.4	1.8 ± 0.4	Thin leaves
IAA 1.0 + kinetin 0.5 (B)	7.5 ± 0.6	3.6 ± 0.5	2.6 ± 0.4
IAA 3.0 + kinetin 0.5 (C)	14.8 ± 2.3	7.8 ± 0.7	5.7 ± 0.5

culture medium, which consisted of Murashige and Skoog[2] macro- and micro-nutrients, vitamins, sucrose (3%, w/v), phytagar (0.3%, w/v) and different combinations of indole acetic acid (IAA) (0.5, 1.0 and 3.0 mg l^{-1}) with kinetin (0.5 mg l^{-1}). The pH of the medium was adjusted to 5.7. Fifteen ml medium was dispensed in 50 ml (2.5 cm diameter) Borosil culture tubes for autoclaving at 1 kg/cm and 120°C for 20 min. Cultures were incubated under 16 h/day photoperiod and 8 h dark period at 28 ± 2°C during the day and 24 ± 2°C during the night at 3000 lx light intensity.

Data on plantlet regeneration, flowering and pod formation were recorded. Ten cultures each in four replications for every treatment were initiated and the data were subjected to statistical analysis.

When curd explants were cultured on MS medium supplemented with different combinations of IAA and kinetin (Table 1), complete plantlets were regenerated within 30 days in treatment B (1.0 mg l^{-1} IAA and 0.5 mg l^{-1} kinetin) and treatment C (3.0 mg l^{-1} IAA and 0.5 mg l^{-1} kinetin). In cultures containing IAA and kinetin each at 0.5 mg l^{-1} (treatment A), only thin leaves were developed and root development was also poor in comparison to the other two treatments.

The phenomenon of complete plantlet regeneration had a bearing on the flowering and pod formation *in vitro* as is evident from the data in Table 2. While flowering initials appeared within 32 days of culture initiation in treatment C along with complete plantlet regeneration, 80 and 45 days were required for initiation of flowering in treatments A and B respectively. Similarly, opening of flowers (40 days) (Figure 1) and complete pod formation (65 days) (Figure 2) were also earlier in treatment C in comparison to treatment B (56 and 95 days, respectively).

By the time complete pod formation was observed in treatments B and C, only flowering initials appeared in treatment A and the medium was also exhausted in culture tubes. Hence, further data on flowering and pod formation are presented with respect to treatments B and C only (Table 3). The number of flowers/plantlets (7.2 ± 0.6) and the length of the flowering stalk (8.6 ± 0.9) were more in treatment C than in treatment B (3.3 ± 0.4 and 4.3 ± 0.4, respectively) after 60 days of

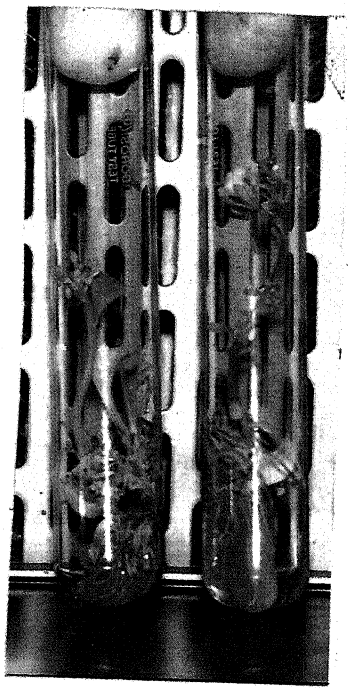


Figure 1. *Left*, appearance of flower initials in treatment B after 45 days of culture initiations; *Right*, opening of flowers in treatment C after 40 days of culture initiation.



Figure 2. Complete pod formation in treatment C after 65 days of culture initiation.

Table 2. Days taken for *in vitro* flowering and pod formation in cauliflower

Treatment	Days taken to		
	Initiation of flowering	Opening of flowers	Complete pod formation
A	80 ± 7.5	—	—
B	45 ± 4.9	56 ± 5.8	95 ± 8.2
C	32 ± 4.3	40 ± 4.8	65 ± 6.2

Table 3. Details of *in vitro* flowering and pod formation in cauliflower

Treatment	No. of flowers/ plantlets	Length of flower stalk (cm)	No. of pods/plantlets	Pod length (cm)	No. of seeds/pod
	(after 60 days)			(after 90 days)	
B	3.3 ± 0.4	4.3 ± 0.4	1.8 ± 0.3	2.1 ± 0.4	—
C	7.2 ± 0.6	8.6 ± 0.9	5.2 ± 0.6	4.8 ± 0.4	4.6 ± 0.5

culture initiation. Similarly, the number of pods/plantlets (5.2 ± 0.6) and the pod length (4.8 ± 0.4) were also more in treatment C than in treatment B (1.8 ± 0.3 and 2.1 ± 0.4, respectively) after 90 days of culture initiation.

After 90 days, shedding of leaves and other floral parts was observed in treatment C and pods were ready to harvest. After harvesting, on an average 4.6 ± 0.5 seeds were counted in treatment C only (Table 3).By this time, pods in treatment B were green andimmature.

The *in vitro* developed flowers were typical of cruciferous family, having 4 sepals, 4 petals, 6 stamens, of which 2 were short, and 2 carpels with superior ovary. The seeds in the fruit (siliqua) were small, globular, smooth and dark brown in colour.

Flowering has also been induced *in vitro* under appropriate conditions through culture of thin cell layers of epidermal tissues excised from flowering stems of tobacco and potato[3]. It is hypothesized that the amount of native gibberellins present in the plant appears to be a deciding factor in whether a plant would remain vegetative or bolt and flower[4]. The effect of gibberellins and other growth retardants on *in vitro* flowering of vegetative apex of Japanese pear has also been reported[5]. In our study, *in vitro* flowering in cauliflower through culture of curd explants appears to be under the control of different hormonal regimes, i.e. 3 mg l^{-1} IAA and 0.5 mg l^{-1} kinentin in MS medium.

Though *in vitro* pollination and fertilization gives a small amount of seed, plant breeders can exploit this phenomenon when this is difficult *in vivo*[6]. Further, the problem of isolation distance is also overcome and the long period for flowering and pod formation under field conditions is considerably cut short this way.

1. Thompson, H. C., and Kelly, W. C., *Vegetable Crops*, Tata McGraw-Hill, New Delhi, 1983, p. 275
2. Murashige, T. and Skoog, F., *Plant Physiol.*, 1962, **15**, 473.
3. Trans Than Van, K., in *Plant Tissue Culture and its Biotechnological Applications* (eds Barz, W., Reienhard, E. and Zenk, M. H.), Springer-Verlag, Berlin, 1977, p. 367.
4. Devlin, R. M. and Witham, F. H., in *Plant Physiology*, CBS Publishers, Delhi, 1986, p. 390.
5. Hignshiuchi, Y., Kido, K., Jchii, T., Nakanishi, T., Kawai, Y. and Ozaki, T., *Scientia Hort.*, 1990, **41**, 223.
6. Pierik, R. L. M., *In Vitro Culture of Higher Plants*, Martinus Nijhoff, Dordrecht, 1987, p.239.

Received 15 February 1994; revised accepted 21 July 1995.

Testosterone biosynthesis in triploid sterile male tilapia

R. Koteeswaran, S. G. Sheela and T. J. Pandian

Department of Genetics, School of Biological Sciences, Madurai Kamaraj University, Madurai 625 021, India

Histochemical localization of strategically important steroid dehydrogenase (3β-HSD) and immunohistochemical studies using testosterone antisera confirmed the active nature of triploid interstitial cells and the presence of 5-ene and 4-ene pathways leading to the biosynthesis of testosterone in the testis of triploid sterile tilapia. Hence, the partial sterility suffered by triploid male tilapia is not due to the lack of steroid hormone involved in spermatogenesis.

TRIPLOIDY has been induced by subjecting freshly fertilized eggs to thermal (heat[1] or cold[2]), pressure[3], electrical[4] or chemical[5] shock; in these shocked eggs the extrusion of second polar body has been suppressed as evidenced by the metaphase-spread preparations. Triploidy is also known to result in complete sterility in females (e.g. rainbow trout[6], tilapia[7]), and partial sterility in males (e.g. rainbow trout[6], tilapia[7]). The partial sterility suffered by males has generally been attributed to endocrine and cytological incompatibilities, although no

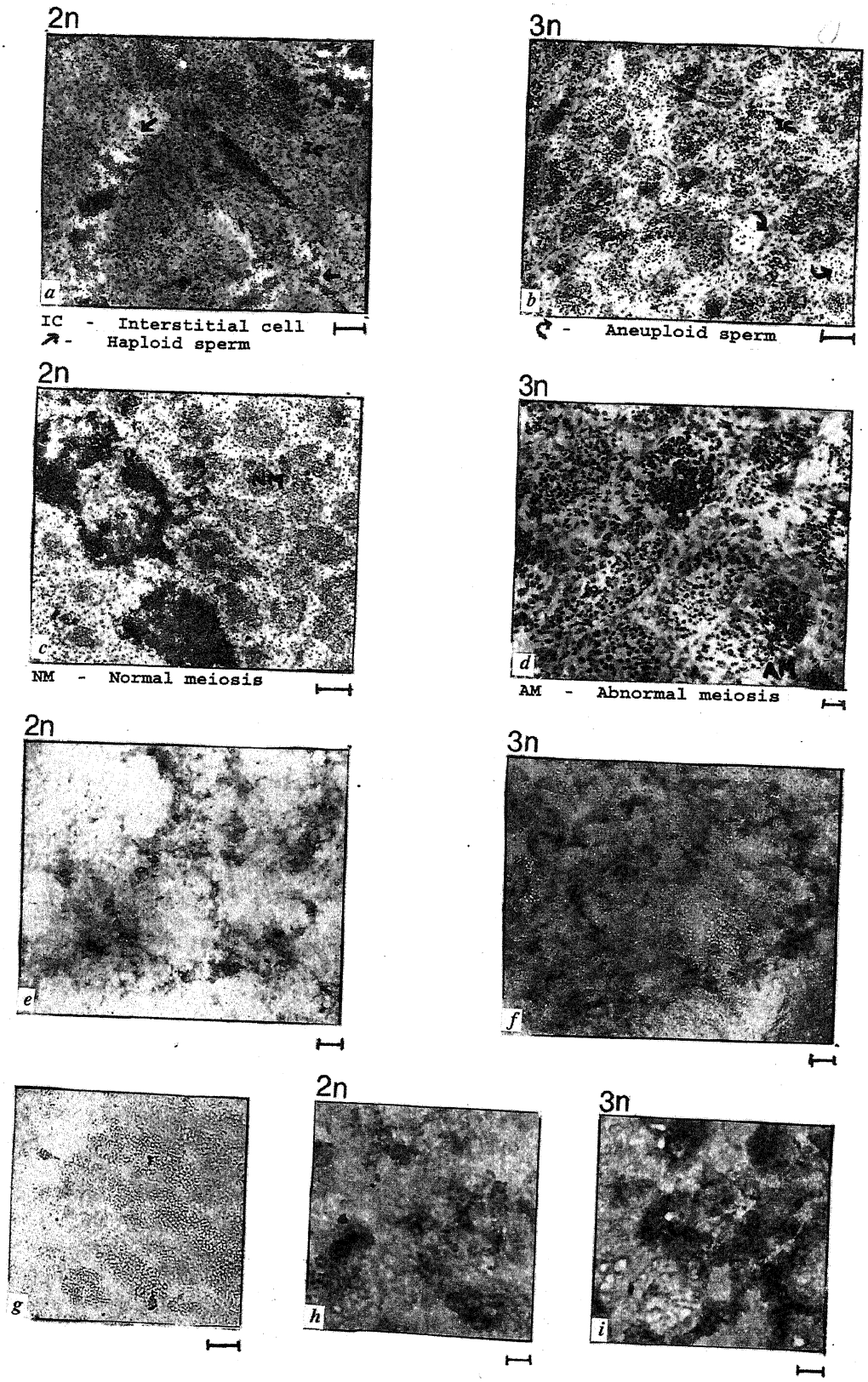


Figure 1 *a–i*. Selected sections of diploid ($2n$) and triploid ($3n$) tilapia. Bar = 5 μm.

direct evidence has so far been reported. Lincoln and Scott[6] were perhaps the first to show the occurrence of testosterone in the partially sterile male triploid *Oncorhynchus mykiss* using radioimmunoassay. This observation has been subsequently confirmed in the same[8] and *Perca flavescens*[9]. This communication reports the occurrence of strategically important hydroxy steroid dehydrogenase, which indicates that the partial sterility suffered by triploid male tilapia is not due to the lack of steroid hormone involved in spermatogenesis.

Triploidy was induced in *Oreochromis mossambicus* following the procedure described by Varadaraj and Pandian[10]. These triploids were reared until they attained sexual maturity. Testes from confirmed triploids were sectioned in a cryotome (International Equipment, USA). H-E staining was done to obtain the structural details. Fresh frozen sections (16 μm) were incubated in an appropriate incubation medium for 2–3 h according to the method of Baillie *et al.*[11]. Pregnenolone was selected as a substrate for histochemical localization of 3β-HSD. The incubation medium contained 0.1 M sodium phosphate buffer (pH 7.2), 0.14 mM pregnenolone, 0.6 mM nitro blue tetrazolium and 7 mM NAD. The sections were fixed in 10% formalin after the completion of reaction. Immunohistochemical staining was done according to the method of Kuhlmann[12]. Sections (16 μm thick) were first neutralized of their endogenous peroxidase activity by treating with 1% hydrogen peroxide in PBS. To reduce nonspecific staining, sections were blocked using normal goat serum and 1% BSA. This was followed by incubation with testosterone antisera (Sigma, USA, in 1:10 dilution of the reconstituted lyophilized stock) for 2 h at room temperature. Finally, sections were treated with 1:100 dilution of HRP-coupled goat antirabbit IGg (Genei, Bangalore) for 1 h and developed in a solution containing 30 mg 4-chloro-1-naphthol, 10 ml methanol, 30 μl of 30% hydrogen peroxide and 40 ml of 1 × PBS for 15 min in dark. After each treatment the sections were washed thoroughly with PBS (pH 7.2) to remove the unbound protein molecules from the sections.

H–E staining of the testis of 5-month-old triploids revealed the presence of numerous interstitial cells as in diploid controls (Figure 1 *a*, *b*). Their active nature was confirmed by lipid and histochemical staining. There were a few irregular-sized spermatozoa in triploid testis, indicating their aneuploid nature arising due to abnormal meiosis (Figure 1 *a*, *d*). In contrast, the diploid sections showed the presence of abundant, regular-sized spermatozoa (Figure 1 *a*, *c*). The 3β-HSD activity was slightly stronger in the triploid interstitial cells (Figure 1 *f*) than that of the diploids (Figure 1 *e*). Control sections incubated in substrate-free medium showed no reaction. Testosterone immunoreactivity was slightly stronger in the triploid sections (Figure 1 *i*) than that in the diploids (Figure 1 *h*). Control sections incubated in substrate-free medium showed no reaction (Figure 1 *g*). Hence, the triploid interstitial cells are active and produce testosterone through the 5-ene and 4-ene pathways.

Hyder *et al.*[13] observed an increase in the plasma testosterone titre in correlation with an enlargement of interstitial cells when human chorionic gonadotropin was administred. Hence, the presence of larger interstitial cells in triploid tilapia might be due to their active involvement in testosterone biosynthesis.

Conversion of pregnenolone to progesterone is a key step in the formation of all biologically active steroids. As 3β-HSD brings about this conversion, its activity has been used to confirm the interstitial cells[14] and is also considered as a measure of steroid biosynthesis by many workers[15,16]. The stronger activity of this enzyme in triploid interstitial cells indicates the presence of highly active 5-ene and 4-ene pathways in triploids. Stronger immunohistochemical staining for testosterone confirms the conclusion.

In conclusion, the triploid interstitial cells and the 5-ene and 4-ene pathways leading to the biosynthesis of testosterone are active. The partial sterility suffered by triploid male tilapia may be attributed to the problems in meiosis, and not due to the lack of steroid hormone involved in spermatogenesis.

1. Kavumpurath, S. and Pandian, T. J., *Aquacult. Fish. Manag.*, 1992, 141–150.
2. Lincoln, R. F., Aulstad, D. and Grammeltvedt, A., *Aquaculture*, 1974, **78**, 287–297.
3. Cassani, J. R. and Caton, W. E., *Aquaculture*, 1986, **55**, 43–50.
4. Teskeredzic, E., Teskeredzic, Z., Donaldson, E. M., McLean, E. and Solar, I., *Aquaculture*, 1993, **116**, 287–294.
5. Johnstone, R., Knott, R. M., McDonald, A. G. and Walsingham, M. V., *Aquaculture*, 1989, **78**, 229–236.
6. Lincoln, R. F. and Scott, A. P., *J. Fish Biol.*, 1984, **25**, 385–392.
7. Pandian, T. J. and Varadaraj, K., *Bull. Aquacult. Assoc. Canada*, 1988, **88**, 134–136.
8. Benfy, T. J., Dye, H. M., Igor Solar, I. and Donaldson, E. M., *Fish Physiol. Biochem.*, 1989, **6**(2), 113–120.
9. Malison, J. A., Procarione, L. S., Held, J. A., Kayes, T. B. and Amundson, C. H., *Aquaculture*, 1993, **111**, 121–133.
10. Varadaraj, K. and Pandian, T. J., *Proc. of the Aquacult. Internatl. Congress*, Vancouver, Canada, 1988, pp. 531–535.
11. Baillie, A. H., Ferguson, M. M. and Hart, D., *Development in Steroid Histochemistry*, Academic Press, New York, 1966.
12. Kuhlmann, W. D., *Immuno Enzyme Techniques in Cytochemistry*, Verlag Chemie, Weinheim, 1984.
13. Hyder, M., Shah, A. V. and Kirschner, M. A., *Endocrinology*, 1970, **87**, 819–822.
14. Jana, N. R. and Battacharya, S., *Zool. Sci.*, 1993, **10**, 489–496.
15. Patra, P. B., Ghosh, A. and Biswas, N. M., *Indian J. Exp. Biol.*, 1976, **14**, 603–605.
16. Kirubakaran, R. and Joy, K. P., *Indian J. Exp. Biol.*, 1988, **26**, 905–907.

ACKNOWLEDGEMENT. Financial support by the Department of Biotechnology, New Delhi, is gratefully acknowledged.

Received 12 June 1995; revised accepted 3 August 1995

Ethnobotany and the Search for New Drugs. (Ciba Foundation Symposium 185, Symposium on Ethnobotany and the Search for New Drugs, Fortaleza, Brazil, 30 November to 2 December 1993). Derek J. Chadwick and Joan Marsh, eds. John Wiley & Sons, Chichester, UK. 1994. pp. 10 + 280. Price: £47.50/$76.00

Ever since the English physician William Withering brought *Digitalis purpurea*, till then part of folk medicine, into Western medicine, plant-based medicines have played a major role in organized healthcare delivery throughout the world. A major step forward in the ascent of plant-based medicines was when, in the last century, Bayer, the German chemical major, decided to synthesize and mass-produce aspirin – a slightly modified form of a naturally occurring compound isolated from willow bark. Aspirin turned out to be the largest sold drug ever! Today more than 25% of prescription drugs issued in North America contain bioactive compounds that are derived from or modelled after plant natural products. Norman Farnsworth of the College of Pharmacy of the University of Illinois at Chicago has identified 119 secondary metabolites isolated from higher plants that are being used globally as drugs. But for awhile, especially after the invention of antibiotics and the development of sophisticated synthetic methods, drug companies relegated the development of plant-based drugs to the backburner. One reason for this temporary disenchantment with plant-based medicines could be the enormous costs involved in testing thousands of plants at random. For instance, the National Cancer Institute of the United States used this approach from the mid-1950s until 1982. The actual yield of this programme is estimated to be one marketable anticancer drug per 8000 plants tested! In the meanwhile drug companies also found the synthetic route – wherein hundreds of compounds are synthesized and tested – to be unaffordable.

It is at this juncture that the enormous potential of ethnobotany came to be recognized. The idea is simple. Instead of testing plants at random, only those plants which are being used by traditional societies to cure their sick will be tested. In other words, one should try to take advantage of folk knowledge. That involves cross-cultural transfer of knowledge. The actual transfer of the plant material from its natural habitat to the drug companies involves transfer of valuable resources. These lead to questions of intellectual property rights, patents and adequate compensation.

Thus, ethnobotany is not mere botany; it is about people, it is about cultures, it is about biodiversity, it is about the possible threats of deforestation, acculturation and wiping out of cultures and natural resources, and it is about multimillion dollar drug business. Above all, it cannot wait. No one knows how long the plants and the cultures which know about their efficacies will survive.

I am sure all these concerns were uppermost in the mind of Paul Alan Cox of Brigham Young University when he suggested the holding of this symposium to the Ciba Foundation. If the resulting book is any indication, the conference must have been organized very well. Some of the leading workers in the field have contributed excellent papers.

The book opens with a brief introduction by Ghillean Prance, Chairman of the Royal Botanical Gardens, Kew. He set the tone for the conference by drawing attention to the interdisciplinary nature of the field, and asking 'What have indigenous people gained from the use of their knowledge by Western culture?' He also drew attention to the problem of translation between the categories (of ailments) adopted by the traditional societies and the categories used in modern medicine. This was followed by 15 presentations; it was rather unfortunate that Richard Evans Schultes, one of the world's leading ethnobotanists, could not make it to the conference. However, his brief but informative paper on Amazonian ethnobotany was circulated and discussed at the conference. Schultes emphasizes the need to intensify ethnobotanical field research to be able to win the race against mindless devastation and destruction of rain forests and acculturation. He ends his paper with the plea that 'all scientists involved make every effort possible to encourage the training of more ethnobotanists and ethnopharmacologists willing to carry on field work with people who still have the knowledge of plants and their properties which they have inherited from generations unnumbered'.

Michael Balick, a student of Schultes, observes that the most valuable contributions of ethnobotany to drug development can come from small and tightly focused programmes, especially with plants used to treat hepatitis, diabetes, diarrhoea, gastrointestinal problems, skin infections, fungi and wounds. In his opinion, it is essential to combine medical and ethnobotanical skills to obtain a proper understanding of plant use in terms of both Western science and traditional beliefs. As we do not know the chemical composition and the full medicinal potential of 99.5% of the plant kingdom, Balick would like drug companies to continue the random approach as well. Indeed, the Institute of Economic Botany, of which Balick is the director, is collaborating with Pfizer in exploring the plant wealth of forests in the United States. Paul Cox, another student of Schultes, has looked at the strengths and limitations of the ethnobotanical approach to drug discovery. He lists more than 50 drugs discovered from ethnobotanical leads. He points out that 72% of every research dollar spent in the Western world is spent on cardiovascular illness, neoplasms, nervous system's disorders and microbial diseases. But very few indigenous cultures have recognized cancer and leukaemia and, therefore, the success rate of ethnobotanical route in the discovery of cancer drugs may not be high. On the other hand, the success rate with regard to gastrointestinal disorders, the most frequent type of disease in Third World countries, may be high.

While emphasizing the need to evaluate plants used as traditional medicines, because more than 3300 million people in the less developed countries use medicinal plants on a regular basis, Norman Farnsworth draws attention to the difficulties encountered in ethnopharmacological work. He warns that unless ethnomedical information is collected in great detail and accurately, it will be no better than random selection of plants followed by targeted biological screening. He pleads for more well-trained ethnobotanists. Farnsworth also gives a brief account of the NAPRALERT database, which provides information on the pharmacological effects of extracts of living

organisms, ethnomedical information on the use of plants, and the biological effects of secondary metabolites.

Walter Lewis and Elvin-Lewis explore the basic, quantitative and experimental research phases of ethnobotany with special reference to the medicinal plants of South America. Search for new drugs cannot succeed unless all the three phases are taken care of. Elaine Elisabetsky and Darrel Posey look at the development of plant-based antiviral compounds, particularly for the treatment of gastrointestinal disorders, based on the experience and practices of Kayapó Indians of Brazil. Another paper based on Brazilian plants, by Craveiro *et al.*, deals with natural products' chemistry. It also describes the Living Pharmacy Programme, under which the local people are taught to cultivate and use medicinal plants correctly. Maurice Iwu, Xavier Lozoya and S. K. Jain have reviewed the current status of African, Mexican and Indian ethnobotanical research, respectively. Iwu has listed several Nigerian plants and their medicinal use; Jain has given lists of Indian plants useful in the treatment of intestinal, joint, liver and skin diseases. Lozoya points out that the lack of clinical studies of plant remedies traditionally used in less developed countries is one reason for the failure to come up with new herbal drugs. Could it be that the Western healers and researchers are threatened by the success of the native epistemology which could, unaided by the tools of Western science, discover and retain the knowledge of plant-based cures for millennia? Does the West suffer from a cultural handicap? In this respect, Asian researchers have an advantage over Western researchers, feels Lozoya.

Pei-Gen Xiao gives an account of ethnopharmacology and new drug development in China. There are 7295 species of plants used medicinally in China, and plant-derived drugs make up 45% of the market. Gordon Cragg and coworkers describe the experience of the US National Cancer Institute (NCI) over the past three decades and more in ethnobotany and drug discovery. A welcome feature is a discussion on intellectual property rights, international collaboration and compensation. NCI collects material from foreign institutions such as the Bogor Herbarium in Indonesia and the Botany Department of the Philippines National Museum.

Two researchers from Shaman Pharmaceuticals, Steven King and Michael Tempesta, describe the company's experience in drug development based on leads obtained from native healers. They have a good team of interviewers assisted by competent translators. Shaman's is one major success story among drug companies using the ethnobotanical route to drug discovery. They already have two drugs, Provir for respiratory viral infections and Virend for herpes, in clinical trials. More are expected. Shaman has a scheme of compensating the native people and communities providing the knowledge.

John Barton, who teaches law relating to high technology at Stanford, discusses the intellectual property rights issues involved in ethnobotany and its use in drug development and manufacture. He pins his hope on nongovernmental organizations to develop a uniform agreement on the relative rights of indigenous peoples and of their governments. Gary Martin looks at the prospects of judicious exploitation of plant resources without harming conservation and biodiversity as well as how ethnobotanists working in the field can help in community development and sustainable plant use. Berlin and Berlin raise questions pertaining to preservation, codification and promotion of traditional cultures and knowledge systems, and returns to the native communities. They narrate their experience in Mexico working with the Highland Mayapeople. The book ends with a brief but perceptive conclusion by Ghillean Prance.

This is an excellent book. It deals with a topic of great current interest and urgency and it has brought together the views and experiences of many of the leading scholars in the field drawn from North America, Europe, Asia, Africa and Latin America. It covers all aspects comprehensively. The conference format has contributed a great deal, as is evidenced by the enormous amount of useful information that has come up in the discussions following each presentation. Indeed, the book would have lost considerably if the discussions were not included. The Ciba Foundation and John Wiley & Sons deserve full marks.

The only flaw in this well-produced book is that some pages shown on the contents page do not match with the beginning pages of the articles.

SUBBIAH ARUNACHALAM

Central Electrochemical Research Institute,
Karaikudi 623 006, India.

Environmental Health Perspective; Supplements: Molecular Mechanisms of Metal Toxicity and Carcinogenicity. National Institute of Health, USA. 1994, vol. 102.

Environment Health Perspective is regularly published by National Institute of Environmental Health Sciences, National Institute of Health, USA. This journal regularly publishes research papers, review articles and sometimes status report in certain areas related to environmental health. It also publishes supplements containing lead articles on research work presented at international meetings.

This volume of *Environmental Health Perspective, Supplements* (Vol. 102, Supplement 3, September 1994) contains selected papers from the Second International Meeting on Molecular Mechanisms of Metal Toxicity and Carcinogenicity held in Madonna di Campiglio, Italy 10–17 January 1993.

The field of metal carcinogenesis and metal toxicology is changing at a rapid pace, so much so that it is now being looked into not only at the organism level but also at the molecular level. Selected papers published in this volume deal with several important molecular aspects of toxicity and carcinogenicity, such as oxidative and free-radical effects, genotoxicity, mechanism of resistance, carcinogenetic process, relationship between carcinogenesis and metallothionein, cellular and molecular aspects of toxicity, etc.

The articles of this volume contain interesting data on chromium-cadmium-, nickel- and arsenic-induced carcinogenesis. Data on chromate-induced DNA strand breaks, Fenton-reagent-induced generation, powerful oxidants, and evidences for radical species as intermediates in cadmium/zinc-

metallothionein-dependent DNA damage *in vitro*, etc., have been discussed. A series of studies in this category revealed a central mechanism that damages to the DNA molecule are caused by oxidants induced by various metals and their salts. The DNA damage is related to mutagenic response and genetic toxicity.

In an interesting article, Wang *et al.* from Norton Nelson Institute of Environmental Medicine and the Kaplan Cancer Center of the New York University Medical Center, New York, reports that an arsenic-resistant Chinese hamster cell line shows increased resistance to toxic concentrations of arsenic after pretreatment with a nontoxic concentration. Such an induced tolerance can be completely inhibited by actinomycin D or cycloheximide. Another group from China reports that human fibroblasts are ten times more susceptible to sodium arsenic than Chinese hamster ovary cells. According to epidemiologic studies, arsenic is associated with increased risk for certain types of human cancers, including epidermoid carcinomas of skin, lung cancers, and possibly liver cancers. However, inoculations of inorganic arsenic have failed to induce tumours in most laboratory animals. The paper by Je-Chang Lee throws light on the differential susceptibility between humans and experimental animals to arsenic, indicating a possible mechanism as to why human exposure to arsenic may cause cancer.

Although cadmium is an environmental carcinogen, suspected to cause lung and prostate cancer in humans, the exact mechanism of cadmium carcinogenesis is not known. A low molecular-weight (6000–7000 Da) protein metallothionein is known to protect cells against metal toxicity. Timothy P. Coogan and colleagues of National Cancer Institute, Maryland, suggest that metallothionein gene activity is quiescent in the ventral prostate of mice. This tissue-specific quiescence of the metallothionein gene has been suggested to determine tissue's susceptibility to cadmium carcinogenesis.

Although metallothionein (MT) plays an important role in cellular resistance to metal toxicity, very little is known about the degradation of metallothionein and tumour of the metals bound to it. In an interesting article by C. D. Klaassen and colleagues of the University of Kansas Medical Center, Kansas, it has been reported that lysosomes might be important in degrading metallothionein, and that metal release is a prerequisite for degradation. The order of sensitivity towards degradation is apo-MT >> Sn MT >> Cd MT.

A unique method for fabrication of ultramicrosensors to study metal movement through cell membranes has been described by T. Malinski and colleagues. A number of papers describe the role of various metals in inducing signal transduction mechanisms. Several papers deal with the effects of organometals on cell signalling, effects of low-level dietary supplementation of organic selenium risks arising out of consumption of polluted mussels for lead poisoning, etc.

In general, this volume contains papers dealing with the most up-to-date knowledge in the field of metal toxicity and carcinogenicity, and also with their cellular and molecular mechanisms. Readers would find it quite useful to update their knowledge.

PRASANTA KUMAR RAY

Bose Institute,
P-1/12, CIT Scheme VIIM,
Calcutta 700 054, India.

Annual Review of Biochemistry 1994. C. Richardson, ed. Annual Reviews Inc., 4139 El Camino Way, Palo Alto, California, USA. Price: USA $49, elswhere $55. Volume 63.

During the past six decades we have discovered time and again that the excitements in the area of biological chemistry are periodically updated in this series, although its undisputed eminence has been constantly challenged. Part of the answer to the success of this series lies in a set of vibrant articles on contemporary themes published year after year. The current issue comprises 28 chapters describing, in principle, a variety of different and diverse aspects under the general rubric of biochemistry. Of particular interest to the ever-increasing number of practising biochemists is the preparatory overview by Osamu Hayaishi. The *Sensei* in an inimitable fashion recounts the events that led to the discovery and characterization of a class of ubiquitous enzymes called the oxygenases. In a fitting tribute, the chapter on lipoxygenase translates his original ideas into much broader issues of human health and disease processes.

The chapters can be roughly divided into four groups. The major section of this book focuses on myriad components involved in cell signalling. The body of any organism (yes, it applies even to the unicellular microbes!) probably cannot function without the network of chemical signals – hormones, growth factors, lipid molecules, neurotransmitters – that communicate with diverse population of cells. Research in these areas over the years have proved to be a fertile ground for gaining insight into the complexities of biology – and of Nobel prizes. There are two chapters, one on nitric oxide and the other on polypeptide toxins that exist in the venoms of predators. The former, which gained notoriety as the molecule of the year 1993 is believed to participate as a major regulator in a variety of cellular processes and is also alleged to have a role in influencing the sexual behaviour. Bredt and Snyder provide an excellent description of all one would like to know about NO, but is afraid to ask! Olivera *et al* emphasize the utility of a class of small polypeptide toxins from marine snails and spiders as probes for elucidating the architecture of calcium channels and as therapeutic agents. This chapter is very clear, well written and is reader-friendly! The section on membrane receptors by Krieger and Herz provides a useful compilation of a variety of different problems in this broad field of multiligand lipoprotein receptors. However, the author's proposition of a 'unifying concept' of ligand recognition, at present, appears to be a dangerous exercise. The framework of the processes in the classical paradigm involving the interaction of G protein coupled receptors with ligands leading to the activation of a variety of cellular processes – via kinases and phosphatases – such as perception of odours, increase in the intracellular vesicle trafficking or change in the electrical activity of brain cells is illustrated in chapters by Strader *et al.*, Johnson *et al.*, and

Nuofer and Balch. There are over 20 odd G proteins which receive and coordinate signals from more than 300 receptors. Despite this complexity, works in several laboratories have linked abnormal G proteins to pituitary and testicular tumours, cholera, leukaemia and other forms of cancer. These three chapters are complementary and anyone with some basic knowledge should be able to follow the flow of the theme. The gods of science lately have smiled fondly on people who took up this messy and elusive area and made it elegantly crisp. Finally, in recent years contributions from the fields of genetics and molecular biology have led to the explosive growth in the number and variety of steroid/thyroid superfamily of intracellular messengers. Tsai and O'Malley have summarized the current understanding of their action, at the molecular level and have identified the issues for the future.

In the second section several chapters discuss various cellular processes and cytological organelles. These include synaptic transmission, intermediate filaments and centrosome. Bennet and Scheller emphasize largely the role of synaptic proteins in vesicular fusion, while Kennedy discusses the molecular events underlying synaptic transmission. Although these two chapters are written independently, the authors make a valiant effort in speedily bringing the oriented reader up to date but the unfamiliar reader may find it essential to consult some textbooks. Intermediate filaments are responsible for a myriad of housekeeping and specialized functions in multicellular eukaryotes. Fuchs and Weber summarize the recently gained exciting insights into the functions of intermediate filaments and their liaison in human disease. In the chapter on centrosome, Kellog *et al.*, discuss the components and their general organization.

In the third group, three separate chapters deal with different and diverse aspects that fall under the general category of energy transduction. Trumpower and Gennis illustrate the way by which prokaryotic and eukaryotic organisms couple the redox energy available from the reduction of molecular oxygen, while Howard and Rees point out that the study of nitrogen fixation has grown exponentially during the past few decades to accommodate not only bioinorganic chemists and spectroscopists but also microbial physiologists. Hartman and Harpel bring us up to date in the continuing saga of engineering plants with catalytically most efficient Rubisco. Although it continues as a paradigm for protein engineering, even the most modest goal of converting the less efficient bacterial enzyme into the somewhat more efficient plant enzyme has not been achieved. However, as discussed, the contributions from the application of atomic genetic and molecular biological tools have clarified the mechanics of energy transduction. Spectroscopic studies have erroneously attributed a role for a covalently bound form of quinone in several enzymatic functions. Although the intimation of quinones as redox cofactors seems to be correct, the article by Klinman and Mu encapsulates the problems and clarify the status of quino-cofactors in biology.

Molecular recognition is the underlying dictum of all chemical and biological processes from protein folding– and catalysis– to genetic recombination. This theme is best illustrated in a set of articles in this group. There are two articles related to HIV, although a part of it was covered in the previous volume. This bias is perhaps understandable in light of its immense importance in public health. Katz and Skalka's chapter describes retroviral enzymes, in general, that also serve as focal points for designing therapeutic drugs. A more direct emphasis is found in the article by Jones and Peterlin on transcription initiation at the HIV-1 promoter with analogies drawn to the eukaryotic systems. The function of DNA polymerase is discussed (Joyce and Steitz) mostly from a structural point of view and the mechanistic connections are less clear. The article on the regulation of DNA replication by Coverly and Laskey is more fascinating than that of the simple information carrier. Of particular interest in this section is the punchiest article on homeodomain proteins (Gehring *et al.*). The authors give the reader an idea of the rewards when chemists and biologists ultimately speak in the same terms in order to understand a complex developmental circuitary system. Over the last decade a combination of genetic, biochemical and structural studies have revealed a new level of mechanistic details in steps underlying the mechanism by which the integrity and fidelity of chromosomes is maintained. The chapters on DNA repair, single-strand DNA-binding proteins and homologous pairing and DNA strand exchange discuss in general the proteins involved and the mechanisms of these processes. Research in these areas is so vibrant that it is discussed almost every year in this series. Indeed, many of the same approaches originally designed for microbial systems have now been successfully applied to higher eukaryotic systems.

I have tried to find faults with this book and in some cases found them. It is regrettable that plant biochemistry, lipids, carbohydrates and methodology have been largely overlooked. Another shortcoming, albeit minor, is the article on Caulobacter differentiation, which belongs, in my opinion, either to the *Annual Reviews of Genetics* or *Annual Reviews of Microbiology*. Despite this, I consider the book to be a highly useful reference on a compendium of topics of diverse nature, containing up-to-date bibliography and appropriate illustrations; consequently, I rate it a shoulder above other series in this league. If the goal(s) of the authors is to inspire the newcomers to the subject, in most cases, it is certainly realized. I am confident that the 'jewel in the crown' of biological chemistry will shine again!

K. MUNIYAPPA

Department of Biochemistry,
Indian Institute of Science,
Bangalore 560 012, India.

Joseph Needham (1900–1995)

There is a saying in Arabic:

> *Maut-ul-Aalim Maut-ul-Aalam,* the death of a scientist is the end of an epoch. This is indeed true for Joseph Needham, whose depth of knowledge and wide-ranging interests are matched by a few, if any.

Joseph Needham was born on 9 December 1900. His father was a physician, and, according to Joseph, a pioneer in pathological histology who later specialized in anaesthesia. His mother was a musician and composer. He had a comfortable childhood, but showed from the beginning a sympathy for working class and when only 13 years old argued for socialism with his father. He earned a degree in natural science and taught biochemistry at Gonville and Callis College, with which he remained associated all his life. He married Dorothy, also a biochemist. Both became Fellows of the Royal Society. However, by 1930 he got interested in history of science, particularly in history of Chinese science. It was probably his contact with Gwen-Djen Lee, who later became his collaborator in the project on Chinese science and whom he married after the death of his wife Dorothy. Two events shaped the direction of his life's work.

The first one was his participation in the Second International Conference on the History of Science and Technology held in London in 1931. Some of the papers presented by the Russian delegation, headed by N. I. Bukharin, which left a deep impression on English participants were those of Boris Hessen on socioeconomic roots of principia, N. I. Vavilov's paper on the origins of the world's agriculture, E. Colman on crisis in mathematical sciences and B. Zavadovsky's paper on physical and biological factors in the process of evolution.

The English side was represented amongst others by J. D. Bernal, J. B. S. Haldane, Hyman Levy, L. Hogben, Benjamin Farrington, Joseph Needham and others. Each one of them was afterwards to work out social, political and cultural and economic dimension of science in their fields of specialization.

The second important event was his appointment as a counsellor, appointed by the British Council in 1942 to extend Anglo-Chinese relations. As a counsellor he travelled extensively in China, and brought out photographs and a sort of commentary on the contemporary developments in the science and technology in China. This was published in 1945 as Chinese Science. Thus began his appreciation of communist China.

There is a photograph, no. 48, in this book with Joseph on a donkey, along with three Chinese on an edge of the desert city of Yuchyachuan. I am told that here his party was attacked by bandits. While the Chinese despaired of their and Joseph's life. Joseph got down from his donkey and did a Scottish dance, seeing that, the leader of the bandit got down from his donkey and did a dance, after which both the parties went their way.

The book gives a bird's eye view of the organisation and development of scientific research of free China, providing historical insights. In discussing these developments his horizon was not limited. For instance, talking of cave temples of Chienfutung he compares them with the caves of Ellora and Ajanta in India.

Joseph had wide-ranging interests on nearly every aspect of human intellectual, social, cultural and political activities, and he comments on these with his erudition. In his own words he was, a 'prowler and explorer among ideas'. He had his views on science, philosophy, religion, arts, history, but what was significant about these views was that he was always modifying them and evolving his ideas about them in the light of new knowledge. For instance, talking of science and religion he says that his earlier description of science was rather too narrow and the description of religion certainly much too neoplatonic, idealistic, pietistic and otherworldly[1].

Joseph was deeply concerned with the rise of Hitler and its consequences on science. He delivered a lecture at Cornell University in 1940 in which he described its consequences on international science. In doing so he drew upon the historical development of science, to show that Nazi's attempt ran counter to the scientific tradition[2]. An interesting feature of the lecture he showed through statistical data the decline of science in Germany as was evident from publication of papers. He also produced data on scientists of established reputation, who were exiled from Germany, Austria, Czechoslovakia and Italy. In most of his articles, lectures and talks, he touched upon historical, social, cultural, political and philosophical features, and also compared the developments with non-European culture areas. His systematic approach to the study of problem is brought out in his essay: 'Limiting Factors in the History of Science'. It covers nearly every possible dimension starting from the relation of the investigator to his environment, co-operation of investigators, techniques and balance between observation, experiment and speculation.

Apart from being a great scholar, he was, what is now being termed, an activist. He was involved in controversies of the period, as those with A. V. Hill and Michael Polyani on the nature of science and role of scientists. He was also involved in a Committee which examined the germ warfare in Korea, and other causes involving injustice, or hegemonistic politics.

He was greatly concerned with one question, one of 'the deepest historical questions that can be raised'. Why did the body of systematized knowledge and theory about nature which we call science not develop also in India or China?[3]

Later, writing a new foreword to the papers presented at the Second International Congress on the History of Science and Technology (1931) to the second edition published in 1971 he wrote; 'With the appearance on the scene of intensive studies of mathematics, science, technology and medicine in the great non-European civilizations, debate is likely to sharpen, for the failure of China and India to give rise to distinctively modern science while being ahead of Europe for 14 previous centuries is going to take some explaining.'

While he concentrated on writing the history of science in China, he encouraged scholars in India, as well as those in Islamic Culture area, to write histories of science in these countries. Writing to the then secretary of the National

Institute of Science of India (now Indian National Science Academy) in 1954 March he wrote, 'I am profoundly convinced (with Dr. Hora) that urgent steps must be taken to put the whole subject of the history of science and technology in India on a proper basis, furthermore that the later history is perhaps the more promising to start with, pending for the historical researches into the datings of texts of antiquity.'

Six years before his death he had organized a conference, which represented his perspective on non-European societies and scientific developments. Explaining the purpose of this Conference on The Historical Dynamics of Oriental Societies, he wrote to the Secretary of Commonwealth Science Council on 15 May 1984: 'The conference is meant to involve a multisided discussion aimed at bringing about greater and more realistic understanding of the specificity of the histories of societies such as China, India, Japan and the Islamic world'.

While Needham worked and published his books on 'Science and Civilization in China', covering nearly all the different branches of science and technology, along with their social and philosophical dimensions, he gave references to the contemporary developments in Asia, as well as in Europe. For instance, in the second volume on the association between nature-mysticism and science he dealt with European linkages, as well as China, the Islamic world and India (pp. 94–99). In his discussions on Sufism and science in the Islamic world he brought out the social features as well. In mentioning the role of Qaramatians he mentions: 'That an alliance of this kind should have existed between the mystical scientists and the organized workers is not the least surprising. Since, as cannot be too often repeated, the great cleavage lay between those who were prepared to engage in manual operations and those who considered them unworthy of a gentleman. Between techniques and magical recipes there was no wide gap' (Vol. II, pp. 96).

He was the first person to recognize the division between natural Sufism and esoteric Sufism. The reason for the latter lay in the power of the governing elite in the suppression of unorthodox beliefs, turning the mystical system into purely religious and unworldly forms. This happened to Sufism in India.

There are many aspects of his personality which are revealed in various incidents of his life. Two or three are worth recalling.

Joseph was the Chairman of a Committee for the development of a university in Sri Lanka, one of the members of this Committee was Mr. Chatterjee, the then Vice Chancellor of the Rajasthan University, Jaipur. Joseph had asked me to meet him in Jaipur. I called on him and he started discussing with me a few points regarding the history of science in India, while we were discussing, the Vice Chancellor came reminding him about the meeting with the Governor. Joseph got up, went to his room and brought a few books for me to go through. The Vice Chancellor, seeing him carrying the load of books remarked, 'There are always servants to carry things for you'. Joseph retorted, 'I never do that sort of thing'.

While Joseph was master of Gonville & Caius College, he invited me to tea. When I went there, there was no light. Thinking that he was not there, I switched on the light and found Joseph deep in thought. I apologized for disturbing him, in his thinking. His reponse was, 'I was not thinking, but was only rearranging my prejudices.'

Another lesson I got from him was when he gave me a book to read on crusades. I had read in my childhood a book on crusades in Urdu, where Muslims were believers and Christians were infidels. Muslims were brave, generous and great fighters while Christians were just the opposite. The book in English painted the picture in the opposite – Muslims were infidels, weak, and not brave!!

As I remember him I am reminded of an Urdu couplet:

'You who are now going away, you would be long remembered.'

A. RAHMAN

Tower A, Flat 30,
Zakir Bagh, M. M. Ali Road,
New Delhi 110 025, India.

Subrahmanyan Chandrasekhar

R. Nityananda

Subrahmanyan Chandrasekhar, the astrophysicist, died on 21 August 1995 at Chicago, where he had worked at the University for the last fifty-eight years. He would have turned eighty-five on the 19 of October this year. His last book, *Newton's Principia for the Common Reader*, had just been published by the Oxford University Press, the product of his dedicated studies over the last several years. He was in the process of preparing for a lecture and thus died as he always lived – working.

His career can be summarized simply enough – an Honours degree in physics at the Presidency College, Madras (1926–1929), a doctorate (1930–1933) followed by a fellowship (1933–1937) at Cambridge University in England, then a move across the Atlantic to the Astronomy Department at the University of Chicago, which was located at that time at Yerkes Observatory, about a hundred miles from the city. He married his college classmate Lalitha, who survives him, in 1936. His positions at the University started with a research associateship and culminated in the Morton D. Hull Distinguished Service Professorship, which he held from 1952 till he relinquished it in 1986 because 'it is better to leave when everybody asks "Why are you leaving?" than to stay while everybody wonders "When is this guy thinking of retiring?"'.

The list of his scientific contributions makes clear why Chandrasekhar was a legend in his own time. In 1930, fresh from his undergraduate degree, he combined the principles of quantum mechanics and special relativity to calculate the pressure of the electrons in white dwarf stars. The result was dramatic – there was a maximum mass, roughly one-and-a-half times that of the sun, above which a sphere of cold matter sustained only by electron pressure would continuously shrink under gravity, never finding any equilibrium configuration. There were indeed earlier papers which had noticed that these principles imply a maximum mass. A year later, Landau wrote a characteristically brief paper which pointed out the existence of the limit, and went on to speculate, presciently about stars with the density of atomic nuclei, and incorrectly about new physics – energy nonconservation – coming into play. Chandrasekhar's work was distinguished both by mathematical thoroughness and a conviction that the result was of fundamental importance. Strong conviction was badly needed, because Arthur Eddington, the Plumian Professor of Astronomy at the University of Cambridge, ridiculed the final result as an absurdity, claiming that one of the starting assumptions was in error. Since Eddington could not pinpoint the error in a manner acceptable to physicists ('I think there should be a law of nature to prevent the star from behaving in this absurd way'), an impasse resulted which is hard to understand so many years later. The personal relationship between Chandrasekhar and Eddington seems to have continued in spite of these events, and he did spend some years in Cambridge after his doctoral degree, though they could scarcely have been comfortable ones. The astronomy community, at least in the United States, seems to have reacted by accepting Chandrasekhar into its midst, without taking any stand on his remarkable result. His own reaction was probably the most remarkable and ultimately fruitful. He wrote his first book on stellar structure, an exposition of both existing knowledge and his new ideas. The style is so lofty and detached that the reader senses no hint of the skepticism and confusion then prevalent amongst astronomers. With this he bid goodbye to the controversy. The stand he had taken on the existence of a maximum mass – known to everyone else as the 'Chandrasekhar limit' – was ultimately vindicated by subsequent theory and observation. The award of the Nobel Prize of Physics in 1983 was primarily for this work.



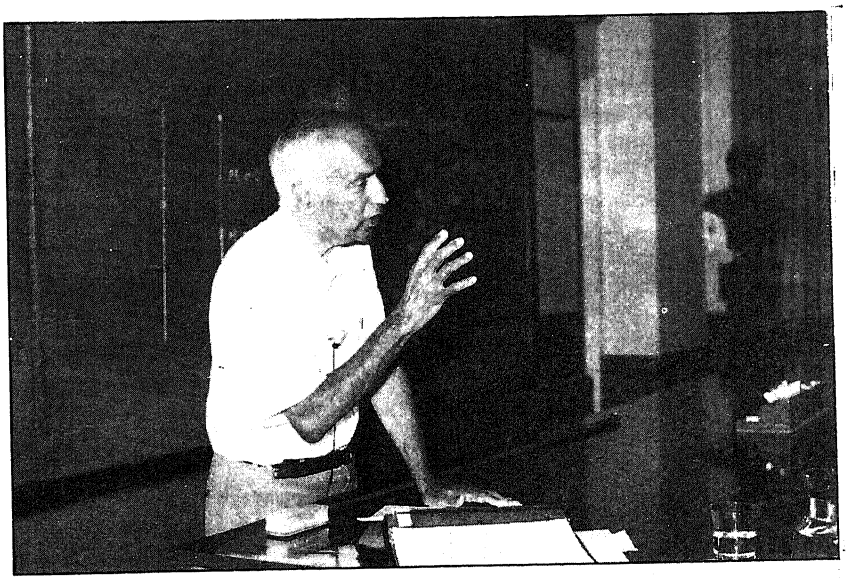
At the Golden Jubilee meeting of the Indian Academy of Sciences.

Around the year 1938, he commenced research in a new area, the dynamics of stars moving in clusters and galaxies under their mutual gravitational attraction. The greatest strength of his work on stellar dynamics was the systematic calculation of the effects of 'encounters' between stars in a cluster – not physical collisions but random gravitational scattering by neighbours. He developed a parallel with the theory describing the Brownian motion of a colloidal particle in a fluid. *En passant*, he wrote the standard survey article on the latter subject as well, in the *Reviews of Modern Physics* of 1943. Generations of theoretical physicists, not necessarily interested in stars, have learnt the basics of Brownian motion from this reference. The deepest consequence of this parallel was that, in spite of the apparent randomness of the scatterings, a star must experience a systematic effect tending to bias its velocity to the mean of its neighbours. His name for this effect – dynamical friction – is as much a part of the vocabulary of astrophysics as the Chandrasekhar limit. Such a process is now known to operate not only in clusters of stars but in much larger scale systems as well, e.g. in merging galaxies.

The work on stellar dynamics ended around 1945 – in this one instance he appears to have written his book, which appeared in 1942, before the picture was

complete. His grand strategy was now established. Five years of deep study and calculation and a book now followed, devoted to radiative transfer – the mathematics underlying the emission, absorption, and scattering of photons in the atmospheres of stars and the pattern of intensity and polarization which emerges to enter the telescopes of astronomers. Today, the problem in its most general and realistic form is attacked by massive computation. Even Chandrasekhar's book begins with a discussion of numerical integration. But after this prosaic start comes a *tour de force* – the nonlinear integrodifferential equations of the theory admit an exact solution in special cases, at least after persuasion by the master craftsman. This calculation gave him the greatest personal satisfaction, stemming not so much from impact as from its elegance. A fruitful physical idea underlies this achievement. While the outcome of the totality of the scattering process in a finite slab of material is usually complicated, the change in the outcome caused by adding one more thin layer can be expressed in a simpler form. This 'invariant embedding' principle also carries with it the 'principle of invariance'. When the slab is semi-infinite, addition of one more layer must reshuffle the individual photons in direction and polarization but leave the overall pattern of the radiation the same.

A systematic study and exposition of hydrodynamic and hydromagnetic instability followed in the next decade, and his book on the subject (which has grown greatly since then) is still valued as a clear and readable introduction, especially to certain topics often used in astrophysics. He made incursions into plasma physics as well. But the next topic to which he turned his attention must have surprised everyone. It was the study of self-gravitating rotating masses of a fluid of uniform density. For one thing, the subject was very old, originating in Newton's 'magisterial' (Chandrasekhar's phase) derivation of the figure of the earth. It had already attracted the attention of mathematicians like Jacobi, Riemann, Dedekind and Poincare, to name but a few. Chandrasekhar could see that all was not well here, and in extensive work with Lebowitz, unified and cleared up many aspects, simplifying the mathematics



Lecturing at the Raman Research Institute

and deriving new results on the way. The achievement was distilled as usual into a book, *Ellipsoidal Figures of Equilibrium*. This severely classical area acquired unexpected relevance with the discovery in 1967 of radio pulsars, which were soon confirmed to be rapidly rotating neutron stars. Since dissipation has a subtle effect on the instabilities of these objects, he was led to include in his studies a form of dissipation caused entirely by an effect of general relativity – the emission of gravitational waves. The electrodynamic analogue is the extra resistance encountered when a charge is accelerated, originating in the loss of energy by radiation. This effect was known from the time of Lorentz, but the gravitational analogue, in spite of work by many people – Einstein, Eddington, Landau and Lifshitz – was riddled with conceptual and mathematical difficulties. A series of papers by Chandrasekhar and his students showed the way by developing the so-called 'Post-Newtonian theory' and deriving, systematically, the gravitational analogue of the Lorentz formula for radiation reaction. In the same period, he also discovered an instability of stars against radial pulsation, again a problem where general relativity made a qualitative change. A well-known astrophysicist has remarked that the work of this period could well have become the content of another classic book. But such a book was never written because by this time the author had now come fully under the spell of Einstein's theory of gravitation, and in particular the black holes, which are an inescapable consequence of that theory (and also his own work on the limiting mass). Unlike on earlier occasions, he was a comparatively late entrant into the area. A whole band of 'relativistic astrophysicists', most of them half his age, were already using pencil, paper, and computer to perturb black holes, scatter all kinds of waves off them, immerse them in external fields, and investigate their properties and stability. The year 1974 brought a heart attack and ultimately major surgery was recommended. He consciously put this off till the fall of 1977 so that he could finish a set of papers on the Kerr solution describing rotating black holes. One of his students has remarked that every complication envisaged by medical textbooks occurred after the operation! It is a measure of his inner strength that he recovered to commence writing *The Mathematical Theory of Black Holes*. This book appeared in 1983 and has a rather different character from its predecessors. It represents his personal efforts to present results in the area (including, of course, many of his own) in a framework which appealed to his own sense of scientific style. Even after all his efforts to find rational and simple ways of doing things, literally hundreds of pages of calculations had to be deposited into archives for scholars of the future. The stability of the Kerr solution could not be fully pinned down. Many things can and have been said of his writings starting from this

period. Understandably, he no longer found the time and energy to go through all the voluminous literature, and perhaps even the proofreading was not as meticulous as before. The problems investigated could border on the exotic – e.g. two black holes connected by a cosmic string! He himself cheerfully described this apparent retreat from the astrophysically relevant as 'downhill, all the way down'. But to focus on these aspects alone would be to miss the basic point – he had done more 'relevant' work than all the critics combined, and was now following where his mathematical and aesthetic instincts led him. What was truly remarkable was the spirit in which all this work was done, which can only be described as youthful exuberance. Consider his paper on the complicated wave equations describing the motion of a spinning electron in the gravitational field of a rotating black hole. From the time of Euler, the mark of a mathematical physicist has been the ability to 'separate' such multidimensional problems by reducing them to one-dimensional components. This is because there is no standard route or guarantee of success. Many people in the field would have been happy to separate the Dirac equation in the Kerr spacetime, but Chandrasekhar started and finished the calculation on the same evening! In this phase, he would fly to Houston for a day or stop over at Crete or Rome for a week to discuss his latest ideas with his friends and collaborators. A young postdoctoral fellow arrived unexpectedly in the same week as when his Nobel Prize was announced. Who else could have got down to work nevertheless and had the first manuscript on colliding gravitational waves almost ready in two weeks? His last book, in which he undertook to read Newton's *Principia* from the point of view of a mathematical physicist rather than a Latin scholar, again bears the strong stamp of being guided by personal taste and satisfaction. His lectures on this subject conveyed vividly the sense of wonder and awe that reading Newton in the original gave him.

There is so much that can be said of his scientific achievements (many have been omitted above!) and style that there is a real danger of missing other contributions. During his watchful and uncompromising editorship, lasting nearly two decades, the *Astrophysical Journal* grew from near-provincial to international stature. His teaching and his supervision of research students were legendary. Even though he had collaborated with the likes of von Neumann (on stellar dynamics) and Fermi (on galactic magnetic fields), he always said his deepest satisfaction came from interactions with students and younger colleagues. Many of these matters, including his austere lifestyle and deep literary and musical interests, have been extensively documented by Kameshwar Wali in a biography entitled *Chandra* (the name by which Chandrasekhar was widely known but which this Indian pen finds too familiar to use comfortably).

A topic of special interest would be his relationship with two countries – India, where he was born and spent his formative years, and the USA, where he grew to world stature. It is often felt that here is a clear case of a person not finding proper conditions and recognition in his own country. His anguish when he felt that a good person in India was not given his due, or with the growth of hierarchy and bureaucracy in the Indian scientific establishment, has been well documented, for example in Wali's book. Yet his relationship with India does not end there. He was a founding Fellow of the Indian Academy of Sciences at the age of 24, and had an offer to work at the Indian Institute of Science after his stay in England. A serious effort was made to bring him to India as the Director of the (then) Kodaikanal Observatory, and he seems to have treated it with equal seriousness before declining. He retained his Indian citizenship till 1953. On visits to Bangalore in the seventies and eighties, he not only willingly and enthusiastically delivered lectures but also spent time talking to students and young researchers, offering comments and reactions in his inimitable style. In these encounters, one could detect no trace of the alienation, cynicism or bitterness that some would ascribe to him. Culturally, he had deep affinities with India, as he has recorded in a little known essay 'on being a foreigner'.

Coming to the American side of the coin, the current wave of adulation launched by the Nobel Prize is not typical of the reactions to him in earlier years, when recognition was slow in coming and hard-won. An authoritative collection of classic and influential papers in astrophysics edited in 1979 managed not to include a single one by him. Perhaps this atmosphere which he encountered in his early years might even have spurred him to new heights of labour and achievement. This is not the place to draw any firm conclusion about the complex feelings that he must have had for both nations, but one can certainly discount the more simplistic view of the matter that is usually propagated. In any case, when one looks at the totality of the work, the manner in which it was carried out, and the impact on colleagues, students, and even people who encountered him just a few times, questions of nationality seem to fade into insignificance. What will remain is the memory and the legacy of a man who never stopped learning, thinking and working throughout a long life, and was driven both by his sense of what was right and what was beautiful to heights that very few will ever attain.

R. NITYANANDA

Raman Research Institute,
Bangalore 560 080, India.

On Ramanujan's bust*

S. Chandrasekhar[§]

Ramanujan was elected a Fellow of the Society in 1918 and he died in 1920, slightly more than a year after his return to India in 1919. The story of how a bust of one who had died in 1920 came to be made 60 years later is of some interest; and I hope that it is proper to tell that story on this occasion.

In his biographical note for Ramanujan (*Proceedings of the Royal Society*, **A99**, xiii–xxix), G. H. Hardy wrote:

'It was his insight into algebraical formulae, transformation of infinite series, and so forth, that was most amazing. On this side most certainly I have never met his equal, and I can compare him only with Euler or Jacobi. He worked, far more than the majority of modern mathematicians, by induction from numerical examples: all his congruence properties of partitions, for example, were discovered in this way. But with his memory, his patience, and his power of calculation he combined a power of generalisation, a feeling for form, and a capacity for rapid modification of his hypotheses, that were often really startling, and made him, in his own peculiar field, without a rival in his day.'

In his lectures on 'Ramanujan' at the Harvard Tercentenary Conference of Arts and Sciences in 1936, Hardy reassessed what he had written in 1921 as follows:

'I do not think now that this extremely strong language is extravagant. It is possible that the great days of formulae are finished, and that Ramanujan ought to have been born 100 years ago; but he was by far the greatest formalist of his time. There have been a good many more important, and I suppose one must say greater, mathematicians than Ramanujan during the last fifty years, but no one who could stand up to him on his own ground. Playing the game of which he knew the rules, he could give any mathematician in the world fifteen.'

In more recent times Ramanujan's reputation has been enhanced considerably. For example, the slight overtone of criticism implied in Hardy's remark 'the great days of formulae are finished' has been countered by 'the days of great formulae are not finished'. With respect to the fall and rise of Ramanujan's reputation, Bruce Berndt (the author of five volumes of 'Ramanujan's Notebooks') has compared him to Johann Sebastian Bach, who remained largely unknown for years after his death in 1750. For Bach, the big turnaround came on 11 March 1829 with Felix Mendelssohn's performance of the *St Matthew Passion*. For Ramanujan, the roughly analogous event was George Andrew's discovery of the 'Lost Notebook' in 1976.

And what is this 'Lost Notebook'?

G. N. Watson had worked on Ramanujan's papers for many years before World War II. When Watson died in 1965, The Royal Society asked J. M. Whittaker (the son of E. T. Whittaker) to write Watson's Biographical Memoir. For that purpose, Whittaker had asked Mrs Watson whether he could examine the papers that Watson might have left in his study. There, in Watson's study, as Whittaker recalled:

'... papers covered the floor of a fair sized room to a depth of about a foot, all jumbled together, and were to be incinerated in a few days. One could only make lucky dips [into the rubble] and, as Watson never threw away anything, the result might be a sheet of mathematics but more probably a receipted bill or a draft of his income tax return for 1923. By an extraordinary stroke of luck one of my dips brought up the Ramanujan material.'

This 'material', of some 87 loose sheets, was part of a batch of papers Dewsbury (the Registrar at the University of Madras during the years 1909–1924) had sent to Hardy in 1923 and that had, somehow, wound up with Watson. After his 'lucky dip', Whittaker passed them on to Robert Rankin (Watson's successor in Birmingham), who in 1968 handed the 87 loose sheets mentioned, together with other unpublished material, to Trinity College, Cambridge.

And there it lay in the Trinity archives without anyone's knowledge till George Andrews rescued it. George Andrews, who had worked on problems related to 'mock theta functions' the subject of Ramanujan's last letter to Hardy a few months before he died, came to Cambridge to explore if any related material was available in the Trinity archives. He was thrilled and excited by what he discovered in the 87 loose sheets deposited in the Archives by Rankin.

Andrews told me, at a later time, that when he presented a paper on the "Lost Notebook" at a meeting of the American Mathematical Society, Dr Olga Taussky-Todd, who was chairing the session, said 'The discovery of the "Lost Notebook" is as sensational a discovery for the mathematicians as a complete draft of a tenth symphony of Beethoven would have been to the musicians'.

Let me continue with what Richard Askey has written in this connection:

'These pages [of the Lost Notebook] are not dated, but from internal evidence they were written late in Ramanujan's life, much of it in his last year. Two thirds of the pages deal with basic hypergeometric series and most of this work is significantly deeper than Ramanujan's earlier work on the same subject. Try to imagine the quality of Ramanujan's mind, one which drove him to work unceasingly while deathly ill, and one great enough to grow deeper while his body became weaker. I stand in awe of his accomplishments; understanding is beyond me. We would admire any mathematician whose life's work was half of what Ramanujan found in the last year of his life while he was dying.'

'Some of Ramanujan's work has one quality which is shared by very little other work. Most mathematics, including some very good work, is predictable. Much of the rest seems inevitable after it is understood, and it would

Reprinted with permission from *Notes and Records of the Royal Society of London*, Volume 49, Number 1, January 1995.
© 1995 The Royal Society.

*Remarks made by the author on the occasion of presenting the bust of Ramanujan to the Royal Society on 11 May 1994.

§ Since deceased

eventually be discovered by someone else. Little of Ramanujan's work seems predictable at first glance and after we understand it there is still a fairly large body of work about which it would be safe to predict that it would not be rediscovered by anyone who has lived in this century. Then there are some of the formulas Ramanujan found that no one understands or can prove. We will probably never understand how Ramanujan found them.'

Let me conclude this part of my story with a quotation from Freeman Dyson:

'The wonderful thing about Ramanujan is that he discovered so much, and yet he left so much more in his garden for other people to discover. For forty-four years I have intermittently come back to Ramanujan's garden; and every time when I come back, I find fresh flowers blooming.'

I now turn to the other strand of the story: how the bust came to be made. It begins with a letter of Hardy's to me.

Hardy was to give a series of 12 lectures on subjects suggested by Ramanujan's life and work at the Harvard Tercentenary Conference of Arts and Sciences in the autumn of 1936. In the spring of that year, Hardy told me that the only photograph of Ramanujan available at that time was the one of him in cap and gown, 'which makes him look ridiculous'. And he asked me whether I would try to secure, on my next visit to India, a better photograph which he might include with the published version of his lectures. It happened that I was in India that same year from July to October. I knew that Mrs Ramanujan was living somewhere in South India, and I tried to find where, at first without success. On the day before my departure for England in October 1936, I traced Mrs Ramanujan to a house in Triplicane, Madras. I went to her house and found her living under extremely modest circumstances. I asked her if she had any photograph of Ramanujan which I might give to Hardy. She told me that the only one she had was the one in the passport which he had secured in London early in 1919. I asked her for the passport and found that the photograph was sufficiently good (even after 17 years) for one to make a good negative and copies. It is this photograph that appears in Hardy's book, *Ramanujan, Twelve Lectures on Subjects Suggested by His Life and Work* (Cambridge University Press, 1940). It is of interest to recall Hardy's reaction to the photograph: 'He looks rather ill (and no doubt, he was very ill): but he looks all over the genius that he was.'

The rest of the story as told by Richard Askey:

'The story of the thread from the "Lost Notebook" to the bust is simple. Andrews has done a lot of very deep work trying to understand what Ramanujan discovered. Eventually the *New York Times* heard about it and interviewed him. *The Hindu* followed with a more extensive interview, and also published an interview with Ramanujan's widow, Janaki Ammal. She lamented the fact that a statue of Ramanujan had never been made, although one had been promised. Andrews sent me copies of these interviews, and after a couple of months my subconscious finally got through to my conscious mind and it was clear that a bust should be made. Since Janaki Ammal was 80, time was important*, so it was up to individuals rather than governments or societies, since institutions move slowly. My first reason for wanting a bust was simple: if Ramanujan's widow wanted one, she should have it. That was the least we could do to show our appreciation of Ramanujan to someone who had been a great help to him.'

'Later I realized there was a second reason, which Janaki Ammal must have realized all along. She knew Ramanujan, and while she did not understand his mathematics, she knew that he was one of the few whose work will last. As long as people do mathematics, some of Ramanujan's work will be appreciated. In Ramanujan's case a more permanent memorial is appropriate; one which can be appreciated by those who do not understand his mathematics should be added to the memorial Ramanujan made for himself with his work.'

From the desire to the accomplishment is a long way.

1. A likeness of Ramanujan had to be found. It was provided by the original of the photograph that is the frontispiece in Hardy's book on Ramanujan.

2. A sculptor who would accept the challenge of making a three-dimensional bust from a two-dimensional photograph. Askey found such a sculptor of distinction in Paul T. Granlund of Gustavus College, Saint Peter, Minnesota, USA.

3. A minimum of four busts had to be commissioned. Askey obtained enough funds by donations from the international community of mathematicians for one bust to be presented to Mrs Ramanujan (which it was in the autumn of 1983). Askey acquired one; and my wife and I two†.

That was how the bust that you see came to be made. And as Askey has said:

'While Granlund does not appreciate Ramanujan's mathematics as those of us who have studied it do, he studied Ramanujan's passport photo deeply, and the results show in the bust. He probably understands some things about Ramanujan that we do not.'

May I now present the bust of Srinivasa Ramanujan to President Sir Michael Atiyah for his gracious acceptance on behalf of the Royal Society.



*Mrs Ramanujan died on 13 April 1994. She was 94.

†Since the original four an additional five have been acquired by other institutions, including the one in the Cambridge Mathematical Faculty Library.

ORISSA BIGYAN ACADEMY

SAMANTA CHANDRA SEKHAR AWARD FOR 1994

(For Oriya Scientists Working Outside Orissa)

The Government of Orissa, Department of Science and Technology, has instituted one Samanta Chandra Sekhar (SCS) Award to be given each year for outstanding research contributions, during the five years preceding the award year, by Oriya scientists working outside the State of Orissa, in any area of (1) physical sciences (2) life sciences or (3) engineering and technology. Persons to be eligible for consideration for this award (a) should be Oriya by birth or domicile, (b) should have passed at least one of the following examinations from any recognized Institution/Board/University in Orissa:

(i) Higher Secondary
(ii) Graduate degree in any branch of science, engineering or technology,
(iii) Postgraduate degree in any branch of science, engineering or technology.

The value of the award is Rs 15,000.00 (Rupees fifteen thousand only) and this is to be awarded through the Orissa Bigyan Academy. Nominations for this award for the Year 1994 are invited from Heads of Research Institutions/Laboratories/Universities/Colleges/Heads of concerned PG Departments in Universities or Colleges in public or private sector, inside/outside the State of Orissa and from SCS Award winners. Nominations from other responsible citizens/government departments may also be considered subject to their acceptance by the Executive Council of the Academy. Five copies of the duly filled-in nomination proforma, along with other particulars stated therein, are to be sent to the Secretary, Orissa Bigyan Academy Type-4R/25, Unit-IX(F), Bhubaneswar 751 007, Orissa (India). The undersigned may please be contacted either by post or in person for obtaining nomination proforma and other particulars. The last date for receiving nominations for this award is 30 December 1995.

Secretary,
Orissa Bigyan Academy
Type-4R/25, Unit-IX(F),
Bhubaneswar 751 007, Tel: 413468.

COUNCIL OF SCIENTIFIC & INDUSTRIAL RESEARCH

NEW DELHI, INDIA.

Advt. No. 8/95

Position: Director, Central Road Research Institute (CRRI), New Delhi

CSIR, the premier agency established by the Govt. of India to undertake scientific and industrial research in the country, is looking for suitable scientist/technologist for the post of Director for its Central Road Research Institute (CRRI), New Delhi. The Institute conducts basic and applied research in the areas of highway engineering and allied subjects, provides know-how for improved and economical methods for construction and maintenance of roads, runways, bridges, etc., and assists the highway user agencies through consultancy, contract research and training. Some of the major R&D programmes of the institute relate to pavement design and performance, road condition monitoring, pavement deterioration modelling, landslide management, hazard mitigation and improved transportation planning for urban areas. In addition, applied research on planning and engineering of rural roads, highway instrumentation, traffic engineering, roads safety and environmental problems form an integral part of the R&D programmes of the Institute.

The post carries the pay scale of Rs 5900–200–7300 plus allowances as admissible to Central Govt. officials. Consultancy is permitted without any ceiling on the annual consultancy remuneration. The appointment to the post will be on contract for a period of six years or superannuation, whichever is earlier.

For further details, please contact the Director-General, CSIR, Rafi Marg, New Delhi 110 001 (India). (Fax: 3710618. Telex: 031-65202 CSIR IN, 031-66147 CSIR IN, Gram: CONSEARCH, NEW DELHI).

SERC Winter School on Databases, Numerical Methods and Computer Modelling in Modern Approach to Petrology (NMCMP-III)

A Winter School on the above subject will be held from 2 to 22 January 1996 at the Department of Earth & Planetary Sciences, University of Allahabad, Allahabad. The NMCMP-III will be directed towards teaching/demonstration of laboratory techniques in experimental petrology, and application of mathematical and computer principles in igneous petrology. Broadly, the course content will include: classification, chemical variation, and structural control of igneous rocks; texture of igneous rocks; phase equilibria relations; petrochemistry of the crust, mantle and core, mantle dynamics; magma: its physical and chemical properties; process of magma generation; diversity of igneous rocks – isotopes and REE-modelling and ore genesis in magmatic systems.

The course is open to university/college teachers, research scholars, scientists of research institutions/professional organizations up to 35 years of age. For details of the course and the application form please write to:

Prof. Alok. K. Gupta,
Course Director (NMCMP-III),
Department of Earth & Planetary Sciences,
University of Allahabad,
14, Chathem Lines,
Allahabad 211 002.

The last date for submission of application is 15 November 1995.

CURRENT SCIENCE

C.V. Raman Avenue, P.B. No. 8001,
Bangalore 560 080.

NEW Display Advertisement Rates*

No. of insertions	Size	Tariff (rupees)					
		Inside pages		Inside cover pages		Back cover page	
		B&W	Colour	B&W	Colour	B&W	Colour
1	Full page	5,000	10,000	7,000	12,000	10,000	15,000
	Half page	3,000	5,000	–	–	–	–
6	Full page	25,000	50,000	35,000	60,000	50,000	75,000
	Half page	15,000	25,000	–	–	–	–
12	Full page	50,000	1,00,000	70,000	1,20,000	1,00,000	1,50,000
	Half page	30,000	50,000	–	–	–	–
24	Full page	90,000	1,80,000	–	–	–	–
	Half page	50,000	90,000	–	–	–	–

SEND PAYMENT BY BANK DRAFT WITH ORDER

*Effective from 1 January 1995

CURRENT SCIENCE

A fortnightly journal of research

Editors: S. Ramaseshan and P. Balaram

Editorial Board

S. Arunachalam, Central Electrochemical Research Institute, Karaikudi 623 006
Asis Datta, School of Life Sciences, Jawaharlal Nehru University, New Delhi 110 067
D. Balasubramanian, Centre for Cellular and Molecular Biology, Hyderabad 500 007
J. Chandrasekhar, Department of Organic Chemistry, Indian Institute of Science, Bangalore 560 012
R. Gadagkar, Centre for Ecological Sciences, Indian Institute of Science, Bangalore 560 012
K. N. Ganeshaiah, Department of Plant Genetics and Breeding, University of Agricultural Sciences, Bangalore 560 065
R. V. Hosur, Tata Institute of Fundamental Research, Bombay 400 005
N. V. Joshi, Centre for Ecological Sciences, Indian Institute of Science, Bangalore 560 012
Jyotirmoy Das, Indian Institute of Chemical Biology, Calcutta 700 032
C. C. Kartha, Division of Cellular and Molecular Cardiology, Sree Chitra Tirunal Institute for Medical Sciences and Technology, Thiruvananthapuram 695 011
S. S. Krishnamurthy, Inorganic and Physical Chemistry Department, Indian Institute of Science, Bangalore 560 012
S. C. Lakhotia, Department of Zoology, Banaras Hindu University, Varanasi 221 005
K. Muralidhar, Department of Zoology, Delhi University, Delhi 110 007
R. Narasimhan, CMC Limited, Bangalore 560 001
K. Neelakantan, ANURAG (Advanced Numerical Research and Analysis Group), Hyderabad 500 258
R. Nityananda, Raman Research Institute, Bangalore 560 080
T. Padmanabhan, Inter-University Centre for Astronomy and Astrophysics, Pune 411 007
B. R. Pai, Propulsion Division, National Aerospace Laboratories, Bangalore 560 017
T. J. Pandian, School of Biological Sciences, Madurai Kamaraj University, Madurai 625 021
Y. S. Rajan, Department of Science and Technology, New Delhi 110 016
M. Ramakrishnan, Geological Survey of India, Calcutta 700 016
A. V. Ramani, T. T. K. Pharma Ltd., Bangalore 560 025
S. Ramaswamy, Department of Physics, Indian Institute of Science, Bangalore 560 012
A. K. Raychaudhuri, Department of Physics, Indian Institute of Science, Bangalore 560 012
S. R. Shetye, Physical Oceanography Division, National Institute of Oceanography, Dona Paula, Goa 403 004
V. Siddhartha, 51 Bharati Nagar, New Delhi 110 003
V. Srinivas, School of Mathematics, Tata Institute of Fundamental Research, Bombay 400 005
G. Srinivasan, Raman Research Institute, Bangalore 560 080
R. Uma Shaanker, Department of Crop Physiology, University of Agricultural Sciences, Bangalore 560 065
K. S. Valdiya, Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore 560 064
G. Venkataraman, Department of Physics, Sri Sathya Sai Institute of Higher Learning, Prasanthi Nilayam 515 134
M. Vidyasagar, Centre for Artificial Intelligence and Robotics, Bangalore 560 001
K. VijayRaghavan, TIFR Centre, Indian Institute of Science Campus, Bangalore 560 012

Current Science, founded in 1932, is published by the Current Science Association in collaboration with the Indian Academy of Sciences. The journal is also intended as a medium for communication and discussion of important issues that concern science and scientific activity. All articles published in *Current Science,* especially editorials, opinions and commentaries, letters and book reviews, are deemed to reflect the individual views of the authors and not the official points of view, either of the Current Science Association or of the Indian Academy of Sciences.

Working Committee, Current Science Association

P. Balaram, Indian Institute of Science, Bangalore
R. Chidambaram, Department of Atomic Energy, Bombay
V. K. Gaur, CSIR Centre for Mathematical Modelling and Computer Simulation, Bangalore
K. Kasturirangan, ISRO Satellite Centre, Bangalore
P. K. Kaw, Institute for Plasma Research, Gandhinagar
N. Kumar, Raman Research Institute, Bangalore (*Vice-President*)
N. V. Madhusudana, Raman Research Institute, Bangalore (*Secretary*)
A. Mani, Field Research Unit, Indian Institute of Tropical Meteorology, Bangalore (*Past-President*)
R. A. Mashelkar, Council of Scientific and Industrial Research, New Delhi
R. Narasimha, Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore
G. Padmanaban, Indian Institute of Science, Bangalore (*Vice-President*)
R. Ramachandran, The Institute of Mathematical Sciences, Madras
P. Rama Rao, Defence Research and Development Organization, New Delhi
S. Ramaseshan, Raman Research Institute, Bangalore (*Treasurer*)
M. S. Valiathan, Manipal Academy of Higher Education, Manipal (*President*)

Editorial Office

Editorial staff

Chandrika Ramesh, G. Madhavan, T. D. Mahabaleswara, C. S. Ravi Kumar, M. S. Venugopal

Circulation and Accounts

Peter Jayaraj, Ranjini Mohan, B. Sethumani, Shanthi Bhaskar, B. K. Shivaramiah

Annual subscription 1995

	Institutional	Personal
India	Rs 350	Rs 100
All other countries (AIR MAIL)	US $175	US $50

Single copies other than special issues: Rs 35/US $15

Editorial Office: CURRENT SCIENCE
C. V. Raman Avenue
P.B. 8001
Bangalore 560 080, India
Telephone: 91-80-3342310
Telex 845-2178 ACAD IN
Fax: 91-80-3346094

Copyright © 1995, Indian Academy of Sciences

INFORMATION FOR CONTRIBUTORS

GENERAL

All manuscripts should be addressed to the Editor, *Current Science*, P. B. No. 8001, C. V. Raman Avenue, Bangalore 560 080. Submission of an article will be held to imply that it has not been previously published and is not under consideration for publication elsewhere; and further, that if accepted, it will not be published elsewhere. *Three copies of contributions of all categories* are required, with a letter of transmittal giving (i) names and complete addresses of the authors and (ii) title of the contribution and the category in which it is submitted (see below).

Current Science is a multidisciplinary journal and therefore research and review papers of general significance that are written clearly and well organized will be given preference. All papers will be first assessed by a Reviewing Editor. Papers found unsuitable in terms of the overall requirements of the journal will be returned to the authors. The others will be sent for detailed review. *Both solicited and unsolicited material will be reviewed.* Authors of these papers will be notified of acceptance, rejection, or need for revision of the paper. Returned papers cannot be resubmitted.

Illustrations and other material to be reproduced from other publications must be properly credited; it is the authors' responsibility to obtain permission for reproduction (copies of letters of permission should be sent).

CATEGORIES OF MANUSCRIPT

General articles (not exceeding 5000 words) discuss current trends in research in a field that will be of interest to readers outside the field; interdisciplinary topics: science policy and administration; or some aspect of the application of science and technology to human needs or the impact of science and technology on society/ecosystems/life. They should include a summary not exceeding 100 words, introductory paragraph(s), brief subheads at appropriate places to point to what follows, illustrations that will help a general reader, and references.

Review articles (not exceeding 5000 words) are expected to survey and discuss recent developments in a field. They should be well focused and organized, and avoid a general, 'textbook' style.

Research articles (not exceeding 4000 words) should report research results of fairly major significance. They should include an abstract not exceeding 100 words, introductory paragraph(s), and brief subheads.

Research communications (not exceeding 2000 words) should contain important findings that are novel and of fairly broad interest. They should include a brief abstract and an introductory paragraph. Text should not be broken up under subheads.

Correspondence includes letters, not exceeding 500 words, that are of general interest to scientists. All letters cannot be published.

Scientific correspondence contains technical comments, including those on articles or communications published in *Current Science* within the previous six months. Letters may be reviewed and edited.

Research news articles are intended to inform nonspecialists about recently published advances or important findings discussed at a meeting. Authors should also send a copy of the paper(s) on which the article is based. Meeting reports should avoid merely listing brief accounts of topics discussed, and must convey to readers the significance of an important advance.

Research accounts articles are intended to be personalized reviews of research from the authors' own laboratory, based on a body of published work. The articles must provide appropriate background to the area in a concise introduction, which should also serve to place the author's work in proper perspective. Articles will normally not exceed 8 to 10 printed pages.

Opinion articles present views on issues related to science and scientific activity. **Commentary** articles should contain expository notes on issues related to science and scientific activity.

Book reviews. Unsolicited reviews will also be considered. Reviews that merely 'list' brief descriptions of the contents cannot be published. Reviews should have 'context' and convey some information about the subject of the book.

Historical commentary and notes inform readers about interesting aspects of personalities or institutions of science or about watershed events in the history/development of science; most are expected to relate to India. Illustrations are welcome. Brief items will also be considered.

MANUSCRIPT PREPARATION

Manuscripts should be typed double-spaced on one side of white bond paper (21 × 28 cm). The pages should be numbered consecutively, starting with the title page and through the text, reference list, tables and figure legends. The **title** should be brief, specific and amenable to indexing. Not more than five **keywords** should be indicated separately; these should be chosen carefully and must not be phrases of several words. **Summary** and **abstract** should not have more than 100 words and should convey the main point of the paper, outline the results and conclusions, and explain the significance of the results.

Text. All papers should have a brief introduction. The text should be intelligible to readers in different disciplines and technical terms should be defined. Tables and figures should be referred to in numerical order. All **symbols** and **abbreviations** must be defined, and used only when absolutely necessary. Superscripts and subscripts and ambiguous characters should be clearly indicated. **Units of measure** should be metric or, preferably, SI. Methods should, as far as possible, be described briefly in appropriate table and figure legends.

Figures. In the case of line drawings one set of originals (without any lettering) is sufficient, with two copies containing lettering. In the case of photographs good prints are required with each copy of the manuscript; photocopies are not acceptable. Line drawings should be roughly twice the final printed size. The correct orientation should be indicated if not clear.

Photomicrographs and other photographs that require it must have a scale bar, which should be defined clearly in the legend. Primary data should be submitted as far as possible (e.g. actual photographs of electrophoretic gels rather than idealized diagrams).

References should be numbered in the order in which they appear, first through the text and then through table and figure legends. The following are examples of ways of writing references.

1. Mukundan, T. and Kishore, K., *Curr. Sci.*, 1991, **60**, 355–362.
2. Constantine. G., in *Biology of Bats* (ed. Wimsatt. W. A), Academic Press, New York. 1970. vol 1, pp. 319–322.

Acknowledgements should be brief. Footnotes are not allowed except to identify the corresponding author if not the first.

Cover photographs. Good photographs (colour or black and white) that pertain to a submitted paper will be considered for use on the cover. Good *prints* and a legend should be submitted with the manuscript. In the case of a colour picture, a transparency will be required for printing if accepted.

PROOFS AND PUBLICATION

Two sets of galley proofs are sent to the corresponding author. A reprint order form accompanies the proofs.

Founded: 1932

ISSN: 0011-3891

CURRENT SCIENCE

Volume 69 Number 7

10 October 1995

563 **In this issue**

CORRESPONDENCE

565 **Assessment of research impact of engineering institutions**
K. L. Chopra

NEWS

566 **Report on the 'International conference on symmetry methods in physics' (ICSMP-95)**
K. Srinivasa Rao and A. Khare

SCIENTIFIC CORRESPONDENCE

569 **Traditional ecological knowledge and sustainable use of natural resources**
R. J. Ranjit Daniels and Jayashree Vencatesan

570 **Endophytes – A crucial issue**
K. R. Sridhar and N. S. Raviraja

GENERAL ARTICLES

572 **Accelerators – Achievements and future projections**
R. Chidambaram and S. Kailas

577 **Recent fertility trends and prospects in India**
K. Srinivasan

REVIEW ARTICLE

587 **Adaptive features of the predatory ladybird beetles in uncertain and competitive habitats**
B. K. Agarwala and S. Bhattacharya

RESEARCH ARTICLES

592 **Radioactivity aspects of Indian coals**
M. P. Rajan, M. A. R. Iyengar and T. V. Ramachandran

596 **Composition and enzyme activities in *Aspergillus flavipes* grown on crude petroleum oil and glucose**
J. Savitha and C. V. Subramanian

RESEARCH COMMUNICATIONS

601 **Liv.52 pretreatment inhibits the radiation-induced lipid peroxidation in mouse liver**
N. G. Ganapathi and Ganesh Chandra Jagetia

604 **An onshore gas well blow-out and its impact observations using satellite data**
K. V. S. Badarinath, A. T. Jeyaseelan, C. S. Jha, Biswadip Gharai and C. B. S. Dutt

607 **Oxygen linkages in neodymium heptamolybdate single crystals – An XPS study**
Sushma Bhat, P. N. Kotru and K. S. Raju

610 **Tectonic evolution of the Central Gujarat plain, western India**
D. M. Maurya, L. S. Chamyal and S. S. Merh

613 **Occurrence and petrogenesis of Loda Pahar trondhjemitic gneiss from Bundelkhand craton, central India: Remnant of an early crust**
Kewal K. Sharma and Abdul Rahman

617 **Hypothalamic γ-glutamyl transpeptidase activity in rats following morphine and naloxone treatment**
P. Sreenivasan and E. Vijayan

619 **Inhibition of biogas production by isobutyric and isovaleric acids and their precursors**
M. K. Sharma and M. M. Simlot

621 **Production of wheat haploids through embryo rescue from wheat × maize crosses**
N. S. Bains, Jaswinder Singh, Ravi and S. S. Gosal

BOOK REVIEWS

624 **Molecular Genetics of Bacterial Pathogenesis,** *reviewed by* J. S. Virdi

625 **Life in Moving Fluids,** *reviewed by* K. S. Gandhi

626 **Annual Review of Neurosciences 1995,** *reviewed by* Gaiti Hasan

627 **Textbook of Molecular Biology,** *reviewed by* S. C. Lakhotia

Corrections

628 **Bose–Einstein condensation in a dilute atomic vapour**
N. Kumar [*Curr. Sci.*, 1995, **69**, 492–493]

628 **Evaluation of efficacy of bricks as geochemical monitor of atmospheric heavy metal pollution**
Rohit Shrivastav, Sanjay Kumar Mathur, Shobhit Shrivastav, M. M. Shrivastav, Sahab Das and Satya Prakash [*Curr. Sci.*, 1995, **69**, 272–275]



COVER. Ground photograph of the ONGC oil well blow-out fire plumes. See page 604.

Indexed in CURRENT CONTENTS/GEOBASE/CHEMICAL ABSTRACTS

In this issue

Dynamics of the Indian population: Changes in fertility patterns

India has emerged as the largest producer of human beings - more additions per year than any other country in the world (including China). The need to cut down this rate has been felt very strongly for more than half a century, and over Rs 5500 crores have been spent on the family planning programme (as of 1990). Paradoxically (or, to be cynical, expectedly), the annual growth rate of 1.97% seen in 1991 is, if anything, marginally higher than the 1.96% seen during 1951-61! Is the situation really as dismal as it appears to be?

A critical analysis of the numbers behind these numbers, described by K. Srinivasan in his article on **page 577** of this issue, however, reveals some very positive aspects. The net growth rate is the difference between the death rate and the birth rate, and though the difference has remained unchanged, both these rates have substantially decreased. Secondly, the fertility rate - average number of children born per woman - has shown a more pronounced decrease, especially in recent years.

The fertility rate is influenced by a number of social, economic, cultural and demographic factors, and this article examines three of the important ones - infant mortality, female literacy and contraceptive use. Multivariate analysis of statewise data across decades consistently shows the female literacy rate to be the dominant factor. Quantitatively, a 20% increase in female literacy is expected to decrease the fertility rate by one child.

What of the future? The article also attempts to project the fertility rates for the next two decades using several alternative scenarios. The good news is that a (highly desirable) fertility rate of 2.0 is quite likely to be attained. See **page 577** for the rationale behind this optimism.

N. V. Joshi

Fungus that feeds on petrol

Brick-biting, blade-chewing and automobile-consuming people have made news too frequently; we have also heard of those who develop a strong addiction to the odour of petrol and kerosene. But it is rare to hear of organisms that 'drink' petrol for their energy needs. J. Savitha and C. V. Subramanian from the Centre for the Advanced Study in Botany at the University of Madras, report of a strain of fungus that does exactly this (**page 596 - 600**). About ten years ago, this group isolated from the effluents of the Madras Oil Refineries, a strain of *Aspergillusflavus* that degrades the hydrocarbons in crude oil. This organism obviously, is potentially useful in cleaning up the ecosystems that are polluted with the crude oil from the industrial effluents or from the oil sleaks in the ocean.

Such organisms, both fungi and bacteria that live in the soil and aquatic systems enriched with crude are being reported ever since the late sixties. Perhaps one of the most interesting among them is the isolation of two strains of *Cladosporium resinae* from the aircraft jet fuel systems by Cofone and his colleagues at the University of Dayton, Ohio, USA. But some of the earlier reports did not clearly establish whether the fungi isolated from the oil-rich soils and marine ecosystems could indeed use and degrade the crude oil on their own. It was suggested that the fungal mycelia might only help in penetrating the insoluble substance such as oil and in increasing the surface area for bacterial attack.

In the recent past, however, these doubts are swept aside and substantial evidence exists to indicate that the fungi indeed consume and degrade the crude oil from the polluted systems. Savitha's group, besides adding another fungus to this list of oil-eating organisms, has also compared for the first time its enzyme activities when grown on glucose with that when grown on crude petroleum oil. Their study shows that *A. flavus* growing on crude oil exhibits 3 to 6-fold enhanced lipid levels with a significantly altered fatty acid profile. They also show that such enhanced lipids in the tissue are brought about by the altered activity in a set of enzymes that probably facilitate the oxidation of alkanes in the crude oil. But what is the significance of such enhanced lipid levels? Is it merely a physiological consequence of 'living' in the oil or is it a physiological adaptation of overcoming the problems of overconsuming the oil? Clearly, the work by Savitha and her group opens up more interesting questions besides being potentially useful in applied microbiology.

K. N. Ganeshaiah

We'll help you get there

If you're selling science books/journals or laboratory products, or wish to recruit scientists in your establishment, then

you're sure to benefit by advertising here.

Because *Current Science* reaches nearly every university, scientific institution and industrial R&D unit in India. What's more, it's read by hundreds of individual subscribers—students, doctoral scientists and professionals in virtually every field of scientific activity in India.

Current Science has the largest circulation among scientific journals in India.

And to give you more impact, we're now bigger, and better.

Write now, or send your copy to

CURRENT SCIENCE

C. V. Raman Avenue, P. B. No. 8001,
Bangalore 560 080

CURRENT SCIENCE

Volume 69 Number 7 10 October 1995

CORRESPONDENCE

Assessment of research impact of engineering institutions

The dictum 'Garbage in: Garbage Out' applies not only to computers but also generally to any creative process. Based on outdated, incomplete and questionable 'scientometric' data of Nagpaul, Prathap (*Curr. Sci.*, 1995, **68**, 869) has drawn some misleading conclusions on assessment of research of major institutions in engineering during 1987–89. Prathap is obviously aware of the fact that 'too little information is a dangerous thing', but all the same he has gone ahead with his analytical exercise.

Let us keep in mind that each engineering institute has its unique features and contributions. We must compare apples with apples and not with oranges. Indian Institute of Science, Bangalore, is primarily a research institute. It has the largest faculty and the largest budget compared to the IITs, and its faculty has little or no teaching. Benaras Hindu University is the largest university in the country with all possible disciplines, including some engineering disciplines. Even the IITs, which have a lot in common (e.g. heavy emphasis on innovative teaching), have many distinctive features. The faculty strength varies from about 300 in IIT Kanpur to over 500 in IIT Delhi. The number of departments/centres varies from about 15 (Kanpur) to 27 (Kharagpur). Even among IITs, a comparison has to be made between comparable departments/centres. And this comparison should include publications, Ph D/M Tech theses, consultancy projects, sponsored projects, industrial interaction, entrepreneural activities, etc. Besides conventional departments/centres, each IIT has some unique departments/centres. For example, Kharagpur has Departments of Mining Engineering, Naval Architecture and Ocean Engineering, Architecture and Planning, and Centres such as Cryogenic Engineering, Rubber Technology, Post-

Table 1. Current Contents on Disc (CCOD)

	PCES 1994	ETD 1994	SBS 1994	Total
IITB	107	88	4	199
IITD	105	164	3	272
IITKAN	100	106	0	206
IITKGP	138	154	2	294
IITM	108	206	2	316
Total	558	718	11	1287
	PCES May 1995	ETD 1995	SBS 1995	Total
IITB	62	67	3	132
IITD	70	103	3	176
IITKAN	71	82	0	153
IITKGP	61	111	2	174
IITM	66	99	2	167
Total	330	462	10	802

Physical, Chemical & Earth Sciences (PCES) 840 Journals
Engineering Technology & Applied Science (ETD) 825 Journals
Social & Behavioural Sciences (SBS) 1360 Journals

Table 2. INSPEC

	1992	1993	1994	Total
IITB	131	142	139	412
IITD	240	152	247	639
IITKAN	131	125	130	386
IITKGP	214	104	205	523
IITM	227	241	223	691
Total	943	764	944	2651

INSPEC is the leading English-language abstracting and indexing service providing abstracts and source information from the world's published literature on all aspects of physics, electronic and computing in approximately 4200 journals, 1000 conferences, books, reports and dissertations, adding over 250,000 records each year.

Table 3. IEEE/IEE publications on DISC (2410 journals)

	1990–May 1995
IITB	82
IITD	153
IITKAN	38
IITKGP	121
IITM	72
Total	466

Harvest Technology, Aquaculture, Biotechnology, Reliability Engineering, etc. Each departments/centre in any IIT has its own R&D culture, which does not necessarily mean publications. Thus, some departments/centres have very few publications but have a very high reputation of their R&D contributions to the industry and other public/private institutions.

Tables 1–3 list comparative statistics of recent publications from the five IITs which are readily available in our Electronic Library and have been obtained from different sources: (1) Current Contents on Disc (CCOD), (2) INSPEC and (3) IEEE/IEE. Each source refers to a group of disciplines, as explained in the captions. Of these groups, the IEEE/IEE publication data are more meaningful for comparison since all IITs have comparable (though not same size) departments in the field. The data in these tables clearly contradict the conclusions of Nagpaul and Prathap and bring out the eminent position of IIT Kharagpur among IITs. The need to do a systematic and in-depth analysis to arrive at 'impact' parameter is clear. Further, such an analysis should take into account other factors such as the number of Ph D's in engineering/technology, number of consultancy and sponsored projects and interactions with industry. Just to give a feeling of the numbers, IIT Kharagpur produces over 100 Ph D's per year, undertakes about 200 projects worth over Rs 1.00 crore as consultancy fee and over Rs 6.00 crore as sponsored R&D projects. Besides, our 30 faculty members have established enterprises, based on their work, in our Technology Park.

Finally, let me emphatically state my faith in critical assessment and accountability of any institution. Indeed, if we do ever hope to see decent standards of R&D in our country, it will be possible only after we have learned to conduct impartial, objective and ruthless peer reviews by nongovernment organizations. Some belated steps taken by UGC and AICTE for accreditation of academic institutions continue to be controversial. It is most unfortunate that our several Academies have failed to appreciate their responsibility to serve the country in this direction. What we need is fiercely independent organizations such as those existing in USA and UK to examine the credentials of each department of an educational institute. We will reach maturity only that day when such an organization will undertake comprehensive analysis to publish publicly 'grades' earned by our academic departments and institutions.

K. L. CHOPRA

Indian Institute of Technology, Kharagpur 721 302, India.

NEWS

Report on the 'International conference on symmetry methods in physics' (ICSMP-95)

The VIIth ICSMP-95, held at the Bogoliubov Laboratory for Theoretical Physics, Joint Institute for Nuclear Research, Dubna, from 10 to 16 July 1995, is the seventh conference in a series, the first five of which were organized by Yu A. Smorodinsky and the sixth was devoted to his memory. This conference was inaugurated by A. N. Sissakian (JINR), and D. V. Shirakov gave some statistical details about the Bogoliubov Laboratory of Theoretical Physics of JINR, where research is carried out in the areas of fields and particles, nuclear theory and theory of condensed matter.

The first (40 min) Invited Talk was by H. D. Doebner (Germany), who spoke on 'Infinite-dimensional symmetries and nonlinear physics'. There were in all 12 Plenary Sessions in which 44 Invited Talks were presented, and 18 Parallel Sessions, in which 85 contributions (of 30 or 20 min duration) were presented. From India, there were three invited speakers (the two authors of this report and M. Lakshmanan of the Bharathidasan University, Tiruchi).

Broadly, the topics covered included the following: dynamical symmetries and integrable models; symmetry and nonlinear phenomena; special functions and group representation theory; quantum algebras and groups; supersymmetry (SUSY) from quantum mechanics to elementary particles; gauge and conformal theories; symmetries in condensed-matter physics and statistical mechanics; geometric methods in quantum mechanics and quantum field theory; quantum optics and coherent states; and mathematical methods.

In a review of the renormalization group (RG) method introduced 40 years ago by Peterman and Stuckelberg, Gell Mann and Low, and N. N. Bogoliubov, D. V. Shirakov (JINR) pointed out that this work led K. Wilson to apply RG method to critical phenomena in the 1970s and that recent analysis identifies the RG with functional self-similarity. In his invited talk 'RG, solution symmetry and Lie group analysis', Shirakov said that Wigner's 1963 Nobel lecture traced the laws of Nature as derived from symmetry and invariance principles and these laws, in turn, explained physical phenomena. The Bogoliubov RG method adds to this scheme Lagrangian parameters and boundary parameters as inputs to the laws of Nature and the physical phenomenon respectively.

P. Winternitz (Canada) in his invited talk on 'Continuous and discrete symmetries of differential-difference equations' showed how the established Lie group analysis of ordinary differential equations (ODE) has been extended to the study of differential–difference equations. The three different approaches to the Lie point symmetries of discrete equations are beginning to converge and similar analysis of q-difference equations

are being made; algorithms, techniques and computerization of these approaches are emerging. M. Lakshmanan (India) in his invited talk on 'Lie symmetries of certain $(2+1)$-dimensional soliton systems' showed how the methodology of similarity reductions applied earlier to $(1+1)$-dimensional problems has been extended to several $(2+1)$-dimensional systems and that all the five cases considered have ∞-dimensional Lie algebras but not all are of the Virasoro type. One of us (A.K.) in an invited talk on 'Singular solutions of KdV and MKdV equations' discussed the construction and properties of the positon and negaton solutions. Just like soliton solutions, these solutions also do not change their shape after scattering and in some cases there is not even a phase change. That the positon solutions are intimately related to 'bound states in continuum' was discovered in 1929 by von Neumann and Wigner. T. A. Weber (USA) in his contribution on 'An extended Gel'fand–Levitan method' presented a study of the resonances of perturbed von Neumann–Wigner potentials.

A. J. Macfarlane's (UK) contribution was on 'Odd-dimensional fermionic phase space and SUSY'. He discussed the algebra of Dirac and Majorana fermions, and showed that fermionic SUSY on a manifold can be constructed. He discussed the case of SUSY with one real supercharge and the problem of a nonrelativistic particle with spin $-1/2$ in a monopole field. S. Duran (Canada) discussed the generalization to 'Fractional SUSY and quantum mechanics', wherein the Hamiltonian is expressed as $H = Q^F$ $(F = 2, 3, \ldots)$, where Q is the supercharge. G. Junker (Germany) discussed Witten's model of SUSY quantum mechanics in his contribution on 'SUSY and its quasiclassical approximation'. The lowest-order quasiclassical approximation is exact for shape-invariant potentials, and for others the resulting approximate spectra are sometimes better than the WKB approximation. M. Zonjil (Czechoslovakia) presented 'Two numerical aspects of SUSY in quantum mechanics', establishing the numerical efficiency of using the Pade approximated superpotentials and the consequences of employing the superpotentials in certain nonvariational treatments like the Hill determinant method. J. Casahorran (Spain) in his contribution on 'Obtaining solvable potentials in the framework of SUSY quantum mechanics' discussed the consequences of SUSY in the context of singular potentials. B. N. Zakhariev (JINR) discussed 'New situation in quantum mechanics (inverse problem and SUSY approach)', wherein he pointed out analytic methods to obtain new potentials whose spectrum and S-matrix are identical to a given potential except for the bound-state energy of one state. L. M. Tomilchik (Belarus) in his contribution on 'SUSY Dirac oscillator: Quantum and classical aspects' discussed the connection between the properties of the Dirac and the classical oscillator in the nonrelativistic and ultrarelativistic limits. A. M. Shirakov (Russia) discussed in his contribution on 'Accidental degeneracy' a situation when there is a degeneracy of levels in the quantum system that does not seem to correspond to any symmetry of the Hamiltonian. As an illustration he offered examples from the spectra of nonrigid spherical molecules.

V. S. Varadarajan (USA) pointed out in his invited talk on 'Quantum semiclassical groups' that corresponding to each method of describing a Lie group a type of deformation is associated and that the Drinfeld–Jimbo deformations are topological algebras. He focused his attention on arithmetic deformations in which real numbers are replaced by p-adic numbers, the real field is replaced by the p-adic field and Hilbert spaces are over p-adic fields. In this scheme, conventional quantum mechanics is obtained as a limit of finite quantum mechanics. In his q-quantum mechanics, p-adic oscillators and p-adic Coulomb problems arise. The hope in q-quantum mechanics is that the divergences of quantum electrodynamics and field theory will be ameliorated. However, deformation of space–time leads to violation of causality, which is troublesome. M. Arik (Turkey) in his contribution on 'Quadratic q-oscillators' gave a formalism which gives as special cases all the well-known quadratic oscillators and, in particular, he constructed homographic, squeezed, square root and chaotic q-oscillators.

G. Rosensteel (USA) in his invited talk on 'Galactic dynamics in the Siegel-half-plane' models the dynamics of rotating galaxies by a Hamiltonian Lax system for which the phase space is a G-manifold with the Lie group G being either the noncompact $\mathrm{Sp}(n, \mathrm{R})$ or a maximal parabolic group $\mathrm{GCM}(n)$. He showed that the dimensions $n = 1, 2, 3$ correspond, respectively, to breathing mode oscillations, planar rotations, and three-dimensional collective motion, and that the homogeneous GCM(3)-manifolds correspond to the Riemann ellipsoids. This work provides a group-theoretic setting for Chandrasekhar's lucid exposition on *Ellipsoidal Figures of Equilibrium* (Yale University Press, 1969). The homogeneous G-manifold $\mathrm{Sp}(n, \mathrm{R})/\mathrm{U}(n)$, where $\mathrm{U}(n)$ is the maximal compact subgroup, was shown to be a classical complex domain diffeomorphic to the Siegal upper half-plane S_n. Equilibrium galactic radii were determined for S_1 systems.

J. P. Draayer (USA) in his contribution on 'Modern applications of SU(3) symmetry and its sp(3, R) extension in nuclear structure physics' pointed out that the SU(3)/sp(6, R) applications started in 1960 for light nuclei $(A < 40)$, extended to the transition region $(40 < A < 100)$ in the 1970s and after a lull in the 1980s picked up interest recently after the success of SU(3) applications to heavy nuclei $(A > 100)$. He discussed some interesting recent work where the origin of pseudospin symmetry in heavy nuclei has been related to chiral symmetry in nuclei. Superdeformed bands have been observed in the $\mathrm{sp}(6, \mathrm{R}) \rightarrow \mathrm{SU}(3) \rightarrow \mathrm{SO_L}(3)$. The pseudo SU(3) scheme has been applied to the study of double β decay of nuclei. D. Benatsos (Greece) in his contribution on 'Quantum algebraic symmetries in nuclei and molecules' discussed super- and hyper-deformations in nuclei and atomic clusters in terms of anisotropic harmonic oscillators with rational ratios of frequencies. I. Pavlichenkov (Russia) in his contribution on 'Symmetrical bifurcation in quantum rotational spectra' considered bifurcation in rotational spectra as a fundamental phenomenon existing in different systems involving different energy scales. Odd nuclei ^{171}Yb and ^{169}Yb show bifurcation, and Coriolis forces tend to align the bands.

K. B. Wolf (Mexico) in his contribution on 'Conformal optical systems' pointed out that the most natural realization of conformal SO(4, 2) symmetry is found in optics. The Maxwell's fish-eye, i.e. the optical system of a transparent medium with spherically symmetric refractive index, had the SO(4, 2) Lie algebra of Poisson brackets. The 'hyperbolic' fish-eye has also been shown to be an optical system possessing conformal SO(4, 2)

symmetry. He pointed out the analogy of the spherical fish-eye to the hydrogen atom problem. A. I. Solomon (UK) discussed the many variations of the successful exact mathematical model of Jaynes–Cummings, which treats a two-level atomic system. In his contribution on 'Jaynes–Cummings variants', he detailed supersymmetric versions of the model and generalizations to the multilevel atomic and multiphoton systems. R. Mc Dermott (UK) in his contribution on 'Coherent and squeezed states parametrized by noncommutative variables' detailed the group-theoretic and Robertson–Schrödinger uncertainty relation approaches to coherent and squeezed states, discussed the properties of the unitary q-displacement Heisenberg–Weyl operator and that of the squeezed vacuum state and its q-generalization. V. L. Derbov (Russia) in his contribution on 'Geometric phases of the generalized polarization of coherent states in quantum optics' showed that the geometric phase provides valuable additional information on the polarization structure of quantum light related to the appropriate dynamical symmetry and suggested experimental measurement of the components of the polarization vector.

One of us (K.S.R.) in an invited talk gave a survey of the 'Quantum theory of angular momentum and generalized (ordinary and basic) hypergeometric functions'. It was shown that sets of hypergeometric functions are necessary and sufficient to account for the symmetries of the $3-j$ and the $6-j$ coefficients, that different ${}_3F_2(1)$ forms for the $3-j$ coefficient known in the literature are derivable from the Van der Waerden form using the transformation theory of hypergeometric functions, that the $9-j$ coefficient's triple sum series due to Alisauskas–Jucys–Bandzaitis can be identified with a triple hypergeometric series, that the $3-j$ and $6-j$ coefficients are related to the Hahn and Racah polynomials, respectively, and that most of these results have been q-generalized. A. Ronveaux (Belgium) in his survey talk on 'Orthogonal polynomials: Recurrence relations for connection and linearization coefficients' presented a simple approach to build recursively the connection coefficients between two families of polynomials and applied it to the classical discrete families (Charlier, Meixner, Krawtchouck, Hahn). M. Noumi (Japan) in his invited talk on 'Reflection equations, quantum-homogeneous space and q-orthogonal polynomials' classified the classical compact symmetry space into ten series and, by the method of reflection equations, he constructed a q-nonsymmetric space (G/K) and, for each (G/K), in seven of ten cases, the resulting zonal spherical functions were expressed in terms of Macdonald polynomials or multivariable Askey–Wilson polynomials of Koornwinder. E. G. Kalnins' (New Zealand) contribution on 'Separation of variables (special functions and modern developments)' detailed how this classical theory is closely related to that of integrable systems and showed how one can systematically classify certain classical mechanical systems via separation-of-variables techniques. R. M. Asherova (Russia) in her contribution on '$U_q(3)$ Weyl coefficients and $SU_q(2)$ Racah coefficients' used the projection operator approach to derive a general analytic formula for transformation brackets between U- and T-bases of $U_q(3)$ irreps. (so-called Weyl coefficients). A resummation technique of q-factorial expressions was used to prove that the Weyl coefficient coincides with the q-Racah coefficient for $su_q(2)$. L. A. Shelepin (Russia) in his contribution on 'Clebsch–Gordan coefficients in coherent and mixed basis' considered a generalized theory of angular momentum including both discrete and continuous parameters and introduced mixed Clebcsh–Gordan coefficients which relate the basis of common $|jm\rangle$ states and coherent states.

The above report presents some of the highlights of the Conference which was well organized and conducted by G. Pogosyan and his dedicated group at JINR. A welcome party, a boat excursion on the picturesque Volga river culminating in a barbecue dinner on an island in Moscow sea, a concert by a Dubna children's choir and a trip to Sergiev Posad, the Russian orthodox centre with its 500-year-old trinity monastery provided adequate relaxation during the week-long deliberations.

The conference concluded with its dedication, by Jelepov of JINR, to the memory of Yu A. Smorodinsky, who along with Pomeranchuk and Migdal was one of the outstanding active theoreticians invited to lecture at JINR soon after it was founded. He recalled with gratitude Smorodinsky's warm personality, broad interests, pioneering work on isospin theory, his new approach to high-energy nucleon–nucleon scattering with a potential including exchange forces, his book *Lectures on Theory of Atomic Nuclei*, with Landau which was translated into French and English (by Dover), and above all the importance he gave to symmetry concepts in physics, which played a big role in the development of this area. It is hoped that this series of conferences would become a permanent feature of JINR, Dubna.

The altered state of the former Soviet Union seems to have one depressing consequence in that, during the past five years, competent people are leaving the country and fewer students have enrolled for higher studies in theoretical physics and pure and applied mathematics – these of the areas in which the people belonging to the former Soviet Union have made deep and profound contributions. We sincerely hope that this is also but a temporary phase in these new nations, which are trying to become economically strong and highly competitive in this materialistic world of ours.

K. Srinivasa Rao, Institute of Mathematical Sciences, Madras, India and **A. Khare**, Institute of Physics, Bhubaneswar, India.

Traditional ecological knowledge and sustainable use of natural resources

A recent publication of the UNESCO discusses 'TEK' – traditional ecological knowledge and its functioning in tropical countries[1]. It identifies three broad categories of TEK, viz.

1) knowledge about specific components/aspects of environment: plants, animals, soils and environmental phenomena which are comparable to systematics,

2) development, evolution and use of appropriate technologies for farming, forestry, hunting, fishing and trapping, and

3) understanding of and intimate relationship with environmental systems as a whole; the most complex and least understood.

TEK has gained much more attention and importance in the field of conservation of biodiversity recently than during the past. Worldwide, a number of government and nongovernment institutions are studying the various dimensions of this natural product both generally and specifically, attempting to build it into the numerous conservation strategies that are being suggested every day. For example, a lot of debate on biodiversity, trade and IPR (Intellectual Property Rights) related issues have stemmed from the overall awareness of TEK; the various round tables on assessment, conservation and sustainable use of genetic resources (prior to the Bahamas Conference of Parties organised by the Stockholm Environment Institute in Latin America, Africa and Asia) and more recently the Seville Strategy of the UNESCO's Man and Biosphere (MAB) programme that was held in Spain (20–25 March 1995). TEK is believed to have been recorded in a variety of ways. They include symbols, phonetics, narratives, rituals, music and dance. Subsistence technology (involving a wide range of natural products), fertility control, settlement patterns, social structure with specific values and norms, acceptance and practice of sorcery, reliance on shamanic powers, forest spirits and supernaturally protected areas (sacred groves) are all considered as tell-tale signs of prevailing TEK in human communities[2].

In the midst of all the emphasis laid on TEK and how best science can acknowledge it, there has always emerged the question 'What do we really know about TEK?' and the statement 'there is a greater need to have information and clearcut examples and case studies'. The best thing at this time would, therefore, be to analyse the Indian TEK that we are familiar with in the context of the conservation and sustainable use of biodiversity and natural resources including air, soil and water.

In India, TEK of medicinal plants and conservation of crop genetic resources in the form of seed banks and locally grown land races have attracted the attention of conservation biologists. Some of these knowledge systems and practices are being systematically documented, especially in the light of developing rewarding systems to the indigenous people who possess them.

Traditional knowledge of macroorganisms in local humans is a fact. For instance, analysis of local names collected of birds in the Uttara Kannada district in the state of Karnataka between 1983 and 1988 suggest that the humans traditionally practised some nomenclatural system as that of modern taxonomy. 169 specific Kannada and Konkani names of birds used by local humans in the district clearly refer to 84 species; a few common and game birds having as many as 7 names. Of these, 38% have a generic identity referring to birds in the same taxonomic family. Another 12% was based on superficial resemblance to unrelated birds, which are rather similar to referring to the coucal as 'crow pheasant' or the chloropsis as 'green bulbul' in English.

Whereas TEK of the various components of the ecosystem and an intimate association with them have been quite convincingly recorded, 'appropriate technologies' for farming, harvesting, etc., that we frequently search for amongst the indigenous people are as yet not fully authenticated although there is little doubt that TEK does include appropriate technologies for harvest of natural resources.

Personal experience suggests that the supportive evidences quoted relating TEK and sustainability of natural resource use by traditional humans are anecdotal and equivocal. For example, some forest dwellers of the Annamalai Hills in south India called *Malasars* are subsisting largely as hunter–gatherers even today. Ethnobotanists who have visited and studied TEK in the use of local plant resources have highlighted the prudent behaviour of these humans with regards harvesting. Tubers of wild plants such as *Dioscorea* are collected only when the plants show signs of maturity. After removing a substantial portion of the tubers, the remaining parts along with the vines are planted back carefully to allow for regeneration. In case where the available tubers are very large, they are shared amongst more members of the community and so on. However, watching the same *Malasars* catching fish in a small river does not provide any evidence of being prudent.

A few men and women together block the flow of the water to start with. An old *saree* is dragged by a pair of *Malasars* along the bottom of the pool and everything that cannot possibly pass through the saree or jump over it is collected, brought to the shore and emptied on the sand. The choice fish are picked up and put inside a small pit dug beside the pool, while the rest including tadpoles and smaller species of fish are just allowed to perish on the shore. Further, the pool remains as such till probably the next rains. Such fishing methods are also practised elsewhere in the tropics of Asia, viz. Northern Thailand[3].

Traditional fishermen catch fish with nets, bamboo traps, pots and even bare hands in many parts of south India. Incidental killing of nontarget organisms is considerable in all these methods. (In our opinion, it is only a magnified form of this practice that we see in modern mechanized fishing wherein 27% of the global harvest is of nontarget species.)

Other examples may be quoted of *Irulas*, hunter–gatherers around Madras. These people randomly set up pit-fall traps and nets in large areas which collect a wide variety of small vertebrate animals. *Irulas* pick up only the choice prey leaving the others in the pits to die. As per the optimal foraging theory, any organism would try to maximize the benefits

derived from the harvested food by way of reducing the energy spent on searching and handling it. A selective mode of harvest of mobile organisms in large areas as that of fish in extensive water involves a greater expenditure of energy than collective harvest and discarding the unwanted. Hence, it is justifiable that humans will naturally resort to the second method of harvest although it involves considerable wastage. In this regard, both traditional and modern approach, as one would expect, would not vary considerably.

Having mentioned a couple of negative examples regarding sustainable harvest of natural resources we may now consider a few positive cases. In south India, the month of July (the onset of southwest monsoon) traditionally used to be a closed season with little harvest of sea fish. Incidentally, this is also the season when most of the choice fish such as seer and pomphrets breed. They come closer to the shore for spawning. Mechanized boats which are being used currently all through the year in all our coastal waters are netting large females and males ready to spawn, considerably reducing the potential for future generations. Although the sea is usually rough during this season, discouraging the use of traditional rafts and boats, one may very well interpret the closed season that used to be observed in the past as a prudent TEK in favour of sustainable harvest of fish.

Instances of local humans using plant poisons to temporarily paralyse fish in pools of freshwater, collecting the select individuals and allowing the rest to 'recover' from the shock are in favour of TEK of sustainable harvest. This is a common practice over most of the south Indian hills. (No effort has, however, been made to record the incidental killings in this method.)

Thus, one can argue back and forth quoting examples which either support or contradict the 'TEK of sustainability' hypotheses. If there are positive examples, there certainly exists an equal number of opposing evidences in this context. Yet the purpose of this article is not to make the readers feel that there is no truth in TEK when it comes to sustainable use of natural resources. This system which has been tried and practised over a few thousand years cannot be ignored altogether. What is needed, however, is an emphasis on more carefully collected and analysed data which is scientifically scrutinized and interpreted. Kenneth Ruddle, writing for UNESCO, says 'study of tribal knowledge in my opinion is not going to tell us more about sustainable harvest but help us understand that they are ecofriendly and can be allowed to exist even within protected areas'. Quoting Ruddle further while closing this brief article on TEK, it is appropriate to state the following. 'The romantic and uncritical espousal of traditional ecological knowledge and management is as extreme and almost as unfortunate as that of dismissing it. Traditional peoples have not lived in some preternatural state of harmony with nature. Some of their abuses of natural resources have been and remain substantial'[4].

1. Inglis, J., *Nat. Res.*, 1994, **30**(1), 2–4.
2. Harp, W., *Nat. Res.*, 1994, **30**(1), 23–27.
3. Kunstadter, P., in *People of the Tropical Rainforest* (eds Denslow, J. S. and Padoch, C.), University of California Press, 1988, pp. 93–110.
4. Ruddle, K., *Nat. Res.*, 1994, **30**(1), 29–37.

R. J. RANJIT DANIELS
JAYASHREE VENCATESAN

M.S. Swaminathan Research Foundation,
3rd Cross Street,
Taramani Institutional Area,
Madras 600 113, India.

Endophytes – A crucial issue

Microorganisms are known to interact with plant surfaces as well as interior tissues. Microbes which live with interior tissues of healthy plants without causing disease symptoms are called 'endophytes', the term synonym of 'mutualists'. Although this concept was confined initially to fungi, later it was expanded to include all microbes which colonize internal plant tissues without causing harm to their hosts[1,2]. In the last decade, major attention was focused on studying the plant interactions with rhizobia and vesicular–arbuscular mycorrhizae (VAM). A few years ago it was understood that a wide range of plants was infected with endophytic fungi. Endophytic fungi have been reported from different plants such as mosses and ferns[3], conifers[4,5], palms[6], and monocots[7]. To date, the ecological role of endophytic fungi is poorly understood. The purpose of this note is to emphasize the gaining importance of endophytes other than rhizobia and VAM.

Endophytes occupy a unique ecological niche and have major influences on plant distribution, ecology, physiology and biochemistry. The studies on the endophytes of grasses, conifers and other woody perennials indicated several beneficial effects. Endophytic infection enhances the resistance of host plants against insects[8]. It was hypothesized that endophytes of conifers and woody perennials decrease palatability for grazing insects and antagonize pathogens[5,9,10]. Thus, to improve their fitness in a given environment, plants might have acquired/accommodated the endophytes.

Endophytes so far isolated were either ascomycetes or their anamorphs and zygomycetes, whereas those belonging to basidiomycetes or their anamorphs or other fungi were rarely reported[3,11]. Several fungi adapted endophytic lifestyle possibly because of the competition of different fungi on suitable substrates. Endophytes are in an advantageous position to colonize the plant parts starting from the initiation of senescence. Endophytes usually bear fruit after the host has died, hence, senescent or dead twigs might be excellent materials to establish anamorph–teleomorph connections.

There is a notion that endophytes have evolved from plant pathogenic fungi. Endophytes may often cause pathological symptoms when the host is under stress[12,13]. These may also behave as latent pathogens, under situations which disrupt the endophyte–plant association (for instance, modern agricultural practices), where endophyte-intolerant host selection will be favoured and transition of endophyte to pathogen might take place[9]. Endophytic, pathogenic and saprophytic behaviour of fungi might be host/environment-dependent. To derive maximum benefits from fungi, plants have to cope with the changing environment to maintain fungi in endophytic rather than pathogenic state, but during senescence of plant parts naturally endophytes switch over to saprophytic phase. Several readjustments have to be made to decide the nature of microorganisms as and when we explore more about endophyte–plant interactions. Some endophytes might behave as pathogens under specific environmental conditions. The decisions taken on microorganisms as plant pathogens or seed pathogens have to be reconsidered with caution. The bacterial or fungal appearance on surface-sterilized excised plant parts in tissue culture can be strongly suspected as endophytes. Plant–endophyte association is an excellent model system which facilitates to understand evolution of plant–microbe symbioses.

Endophytes are known to infect different plant parts either by vertical transmission via host seeds/vegetative propagules or by horizontal transmission. It is now known that climatic conditions greatly influence fungal spore germination and ultimate infection. The increase in rainfall increases the level of endophytic infection. The deposition of surface runoff along the midrib/lamina of leaves favour fungal spores to collect and subsequently to infect heavily[14]. It is also known that aquatic roots of riparian vegetation are colonized by aquatic hyphomycetes, and it is suspected that the selection pressure might have been induced by scarcity of substrates[15,16].

Many plant secondary metabolites have structural and qualitative similarities with endophytic allelochemicals and fungal toxins[17]. The possibility of endophytes acting as vehicles of horizontal gene transfer between the genomes of fungi and plants cannot be ruled out, if the endophytes are not the allelochemical producer. Interestingly, many grass endophytic fungi are known to produce alkaloids which act as insect repellents. Several grasses which possess endophytes (infected by *Acremonium* spp.) are toxic to livestock. Surprisingly, on grazing some grasses cattle, horse and sheep get intoxicated and that is why animals avoid grazing such stuff[8]. It is now known that the intoxicating effect is due to the metabolites of endophytic fungi[18]. The properties of plant parts in curing human/animal ailments and as intoxicants and toxins might also be influenced by the endophytic partner. Endophytic fungi are now recognized as a potential source of new secondary metabolites[19]. Endophytes are known to produce compounds active against human and plant pathogenic bacteria[20–23]. Besides these, endophytes including VAM produce hormones such as indole-3-acetic acid (IAA) and cytokinins which stimulate growth and flowering of plants[24,25]. Endophytic niche is considered as the potential source of microbes which produce a wide variety of bioactive metabolites. It has become a new source of biochemicals like those from soil, freshwater and marine ecosystems.

Dreyfuss and Chapela[26] estimated that over 1.3 million endophytes are awaiting discovery. The study of endophyte distribution, biodiversity and their biochemical characteristics are of immense importance in plant biology to understand and to improve plant fitness; to develop safe biocontrol chemicals; to produce environment-friendly chemicals in industrial scale. The distribution and fitness of plants in a given environment; physiological performance of plants and biochemical features of plant tissues have been greatly influenced by endophytes. There are many benefits to be derived and several questions yet to be answered regarding endophyte–plant interactions.

1. Petrini, O. and Carroll, G. C., *Can. J. Bot.*, 1981, **59**, 629–636.
2. Petrini, O., in *Microbial Ecology of Leaves* (eds Andrews, J. and Hirano, S.), Springer, New York, 1991, pp. 179–197.
3. Petrini, O., in *Microbiology of the Phyllosphere* (eds Fokkema, N. J. and Van den Heuvel, J.), Cambridge University Press, Cambridge, 1986, pp. 175–187.
4. Bernstein, M. E. and Carroll, G. C., *Can. J. Bot.*, 1977, **55**, 644–653.
5. Carroll, G. C. and Carroll, F. E., *Can. J. Bot.*, 1978, **56**, 3034–3043.
6. Rodrigues, K. F. and Samuels, G. J., *Mycol. Res.*, 1990, **94**, 827–830.
7. Fisher, P. J. and Petrini, O., *New Phytol.*, 1992, **120**, 137–143.
8. Clay, K., *Ecology*, 1988, **69**, 10–16.
9. Carroll, G. C., *Ecology*, 1988, **69**, 2–9.
10. Clark, C. L., Miller, J. D. and Whitney, N. J., *Mycol. Res.*, 1989, **93**, 508–512.
11. Petrini, O., Fisher, P. J. and Petrini, L. E., *Sydowia*, 1992, **44**, 282–293.
12. Miller, C. S., in *Microbial Ecology of the Phylloplane* (ed. Blackman, J. P.), Academic Press, London, 1980, pp. 185–209.
13. Andrews, J. H., Hecht, E. P. and Bashirian, S., *Can. J. Bot.*, 1985, **60**, 1216–1221.
14. Wilson, D. and Carroll, G. C., *Mycologia*, 1994, **86**, 635–647.
15. Fisher, P. J., Petrini, O. and Webster, J., *Mycol. Res.*, 1991, **95**, 543–547.
16. Sridhar, K. R. and Baerlocher, F., *Mycol. Res.*, 1992, **96**, 305–308.
17. Wilson, D., *Oikos*, 1993, **68**, 379–384.
18. Bruchl, G. W., Kaiser, W. J. and Klein, R. E., *Mycologia*, 1994, **86**, 773–776.
19. Bills, G. F. and Polishook, J. D., *Sydowia*, 1992, **44**, 1–12.
20. Dreyfuss, M. M., *Sydowia*, 1986, **39**, 22–36.
21. Fisher, P. J., Anson, A. E. and Petrini, O., *Bot. Helv.*, 1984, **94**, 249–253.
22. Fisher, P. J., Anson, A. E. and Petrini, O., *Trans. Br. Mycol. Soc.*, 1984, **83**, 145–148.
23. Fisher, P. J., Anson, A. E. and Petrini, O., *Trans. Br. Mycol. Soc.*, 1986, **86**, 153–156.
24. Pugh, G. J. F., in *Seed Ecology* (ed. Heydecker, W.), Butterworth, London, 1972, pp. 337–345.
25. Allen, M. R., Moore, T. S. and Christensen, M., *Can. J. Bot.*, 1980, **58**, 371–374.
26. Dreyfuss, M. M. and Chapela, I. H., in *Discovery of Novel Natural Products with Therapeutic Potential*, Butterworth, London, 1992, Biotechnology Series.

K. R. SRIDHAR
N. S. RAVIRAJA

Department of Biosciences,
Mangalore University,
Mangalagangotri,
Mangalore 574 199, India.

Accelerators – Achievements and future projections*

R. Chidambaram and S. Kailas

Charged particle accelerators are being used not only for basic research but also for various applications. Accelerator activity started in India almost four decades back and we have now the capability and expertise to design, build and operate many types of accelerators which are required. International collaboration has been pursued in programmes involving large accelerators. In the present article, a brief description of the Indian Accelerator Programme, its achievements and future projections, is given.

CHARGED-particle accelerators are being increasingly used, both directly and indirectly, for research in many frontier areas of science. They are also indispensable for varied applications ranging from materials science to medicine and, more recently, even for radioactive-waste transmutation.

As India has achieved, over the years, the distinction of being able to design, build and operate many types of accelerators which are required, it is timely to take stock of the Indian Accelerator Programme. Just as the design and construction of a reactor involves competence in several areas and integration of several subsystems, so does the design and construction of an advanced accelerator, and the units of the Department of Atomic Energy (DAE) are uniquely equipped to undertake this task. Further, with increasing interest in indigenous accelerators, both for research and for applications, and possibilities of increased participation in international projects, both in construction and for research, it is useful to review our achievements and attempt to define future possibilities in this area. It is inevitable that this paper emphasizes the work done in the DAE – considering the affiliation of the authors – but we have tried to include, to the best of our knowledge, the contributions to this field from the university system.

Historical development up to the eighties

The work related to various types of accelerators began in India in the forties and the fifties, though in a modest way. In the sixties India had a reasonably large size accelerator in the form of the 5.5 MV Van de Graaff accelerator at the Bhabha Atomic Research Centre (BARC), purchased from High Voltage Engineering Corporation, USA. The accelerator-building activity took a quantum jump with the indigenous development and commissioning of an 88 in. variable energy cyclotron at Calcutta (VECC) in the seventies (see Box 1). Around the same time a 2 MV tandem Van de Graaff accelerator was made operational fully through indigenous efforts at Trombay. In the eighties, to catch up with the ongoing research in the important area of heavy ions, we embarked on a medium-energy heavy-ion facility in Bombay (see Box 2). Though the 14 MV pelletron, the first of its kind in India, was imported from National Electrostatics Corporation (NEC), USA, several subsystems such as the 25 m high, 5.5 m wide accelerator tank to contain the SF_6 gas at high pressures, the four storage tanks for SF_6 gas, the beam lines and the experimental facilities were all fabricated and set up indigenously. This joint BARC–TIFR facility located at the Tata Institute of Fundamental Research (TIFR) campus has now been operating for more than five years and in this period with the use of this facility considerable research work has been carried out in the field of heavy-ion nuclear reactions and in interdisciplinary areas.

Growth in the nineties

In the year 1990, the second pelletron, a 16 MV NEC machine was made operational in Delhi at the Nuclear Science Centre (NSC) of the University Grants Commission (see Box 3). To augment the scope of accelerators for areas other than nuclear physics, a 3 MV NEC pelletron was commissioned in the Institute of Physics, Bhubaneswar, in 1991. Work has already started in converting the 5.5 MV Van de Graaff accelerator at BARC into a folded 7 MV tandem accelerator. The Centre for Advanced Technology (CAT) at Indore (see Box 4) was started a few years back for development of technologies associated with accelerators, lasers and related systems. The first Indian synchrotron radiation source, based on a 450 MeV electron storage ring, Indus-1 (see Figure B4), is already at an advanced stage of construction at CAT. This facility when completed will consist of a 20 MV microtron, followed by a 700 MeV synchrotron and the storage ring. This storage ring will

*Based on the XV Prof. Balakrishnan Memorial lecture of the Indian Vacuum Society given on 14 October 1994 at Bombay by Dr R. Chidambaram.

R. Chidambarm and S. Kailas are in Bhabha Atomic Research Centre, Bombay 400 085, India.

provide soft X-ray and vacuum ultraviolet radiation over a wide spectrum. In the next phase, a 2 GeV storage ring Indus-2 (see Figure B4) to deliver harder X-rays through dipole magnets and wigglers will be built.

Other developments related to the accelerators described above include the work started for the building of superconducting LINAC boosters to enhance the heavy-ion energies from the present pelletrons. At VECC, a versatile electron cyclotron resonance (ECR) ion source has been built and will soon be used as an injector to the present cyclotron to deliver high-energy heavy-ions for experiments from this facility. Plans are also under way to build a superconducting cyclotron at Calcutta. At BARC, work is going on in completing an REQ accelerator which, in principle, can also serve as an injector for a bigger accelerator. There are also plans to build low-energy electron accelerators primarily for medical and industrial applications both in CAT and in BARC. CAT is also building a 20 MeV proton Linac which will consist of a 750 keV RFQ operating at 202 MHz and a drift tube Linac also operating at 202 MHz.

Box 1. Variable Energy Cyclotron Centre, Calcutta

The cyclotron at Calcutta became operational in the late seventies and has been delivering beams of alpha particles for not only nuclear physics investigations but also for other areas of science. Recently, both protons and deuterons of different energies have become available to the user community. Beams of high resolution (about 70 keV) are available for research. In the near future it is expected to reach a value of 10–15 keV. When the ECR ion source is coupled to the present machine, a large variety of medium- to high-energy heavy-ions will be available from VECC. A recently completed experiment utilizing 15–20 MeV/A alpha particles interacting with aluminium has given interesting answers to the production mechanism of fragments heavier than the projectile. From systematic studies of break-up of ^{6}Li using alpha particles from VECC, important results of relevance to astrophysics have been obtained (Figure B1). Another area of research is the study of properties of nuclei far from the line of beta stability, an emerging area of great potential. For this purpose, an isotope separator on-line (ISOL) facility has been developed.

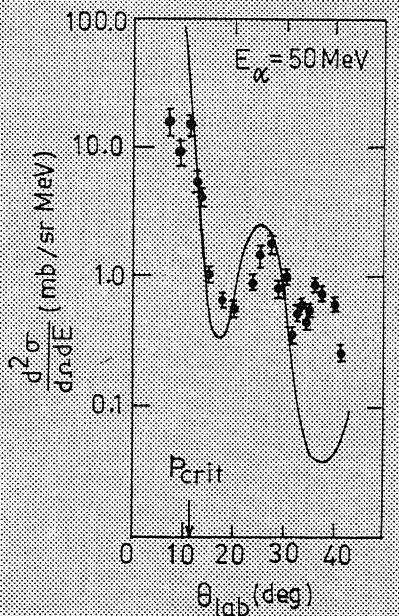


Figure B1. Alpha particle-induced, $^6Li \rightarrow \alpha + d$, target breakup reaction at VECC, Calcutta. The finding of a critical momentum transfer – p_{crit} = 117 MeV/C – which signifies the dominance of the Coulomb breakup, is important in the context of nuclear astrophysics.

Box 2. Pelletron in Bombay

The commissioning of the medium-energy heavy-ion accelerator facility, 14 MV pelletron, in Bombay six years ago has enabled the nuclear physicists in this country for the first time to explore the nucleus in many previously inaccessible regions of excitation, spins and configurations. In order to carry out research in several possible frontier areas, a number of facilities have been set up which include: (a) A general-purpose 1 m diameter scattering chamber of capacity 500 l, (b) Large-area position-sensitive ionization chamber, (c) Array of NaI, BGO, Phoswich and HpGe detectors, (d) Fast electronics and PC/transputer-based data acquisition systems.

Amongst the many interesting results which have come out of this facility, the notable ones are establishing the fission time scales through the use of 'neutron clock' and the experimental verification of the preequilibrium fission mechanism by the BARC group from a series of fission fragment angular distribution measurements (Figure B2). From recent investigations related to quasimolecular resonances in light compound nuclei, it has been possible to establish superdeformed structures in shapes of nuclei with major to minor axis ratio between 2 and 3.

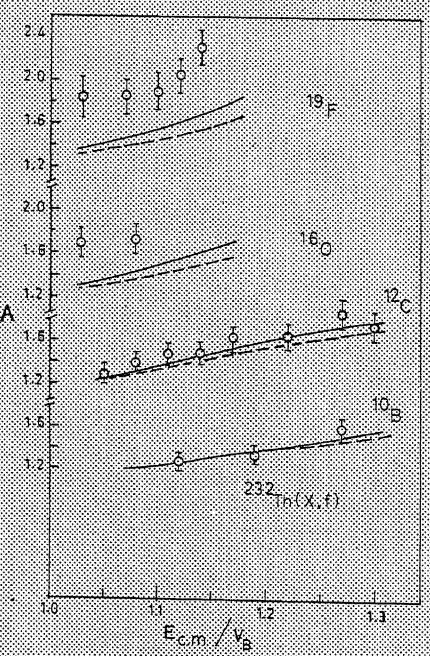


Figure B2. Fission-fragment anisotropies for ^{10}B, ^{12}C, ^{16}O and ^{19}F + ^{232}Th systems measured using the Bombay pelletron. While the systems ^{10}B and ^{12}C + ^{232}Th exhibit 'normal' anisotropies, the ^{16}O and ^{19}F + ^{232}Th systems exhibit 'anomalous' anisotropies. This observation is consistent with the expectations of pre-equilibrium fission model proposed by the Trombay group earlier.

Box 3. Pelletron in Delhi

A heavy-ion facility with a 16 MV terminal has been operational at NSC, Delhi, for the last few years. The NSC user family has members from more than 30 universities, 3 IITs, IISc, besides national labs like NPL, BARC, VECC, IOP, SINP and TIFR. Some of the major experimental facilities which are available in this centre are: (1) The heavy-ion recoil analyser (HIRA) for detection of slow-moving heavy recoils, with a mass resolution of better than 1/200. (2) Gamma detector array consisting of a set of HpGe detectors with anti-Compton shields for high-spin spectroscopy. Additional array of BGO detectors is provided for total energy measurement and it also serves as a multiplicity filter. (3) 1.5 m diameter scattering chamber. Using these facilities a number of experiments have been performed related to subbarrier fusion phenomena and high-spin spectroscopy (Figure B3).

To augment the energies available from these facilities, both at Bombay and Delhi, LINAC boosters based on superconducting technologies are being built. While the one at Bombay will be based on Pb, the one at NSC will be Nb-based. The NSC is collaborating with Argonne National Lab, USA, on this LINAC project.

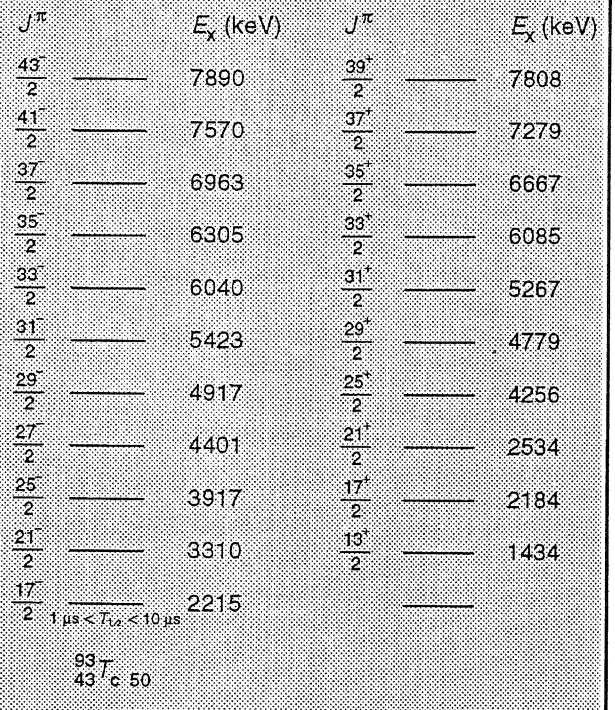


Figure B3. High spin states in ^{93}Tc. Level scheme for ^{93}Tc showing excitation energies and spins (parities), populated through ^{66}Zn (^{31}P, 2p2n) ^{93}Tc reaction at the Delhi pelletron accelerator facility.

A Centre for Compositional Characterization of Materials (CCCM) is being set up by BARC at Hyderabad and some of the analytical techniques used there will employ a 3 MV tandem accelerator. All these go to show that we have in this country several strong groups in the accelerator field and also the necessary infrastructure to design and build indigenously new generations of accelerators and related experimental facilities.

The various accelerator facilities in the country are collectively presented in what can be called the accelerator map of India (Figure 1).

International collaboration

It is very clear that no single country in the world can now afford to build and operate major accelerator facilities and related detector set-ups. International collaborations have become necessary. Our country has been collaborating with major international accelerator facilities in Europe and USA. Our contributions are both in the area of accelerator building and in the development of large-size detectors. The CAT is collaborating with CERN, Europe, both in LEP (large electron–positron collider) and PS (proton synchrotron)/SPS (super proton synchrotron) accelerator upgrades. Titanium sublimation pump bodies have been exported to CERN. It has also supplied other vacuum components and corrector magnets in substantial numbers. A group from BARC and VECC has developed the control software for accelerating Pb ions at SPS facility.

The high-energy physics group from TIFR has been actively collaborating with CERN for a number of years. For the L3 experiment, the TIFR group has supplied 1100 brass proportional chambers, SS housings, precision machined Cu–Te pieces and also a few thousand PCBs. The TIFR group has also contributed to the software and physics aspects of this joint programme, in particular the precision tests of the standard model of particle physics and a search for one of its cornerstones: the Higgs particle. Three major achievements of this experiment are: (1) the number of neutrino species has been established to be three; (2) precision measurement of the electroweak mixing angle; and (3) estimate of the top quark mass.

With the recent availability of relativistic heavy ions from CERN at SPS, the group from VECC has been collaborating with CERN on experiments related to the experimental search for the quark gluon plasma (QGP). The VECC group has built a photon multiplicity detector consisting of 55,000 pads of small plastic scintillators (20 mm × 20 mm) for efficiently detecting the photons as signals of QGP. The optical fibres are inserted diagonally in each pad to pick up the light signal. One successful run has already been completed with a sulphur beam of 200 GeV/nucleon energy using the CERN SPS. Recent analysis of the photon data has perhaps given the first hint of the QGP, entirely from the photon spectrum. Besides VECC, several scientists and students from Institute of Physics, Bhubaneswar, and Indian universities – Chandigarh, Jammu and Jaipur – are also collaborating in this international venture.

Recently, BARC has also entered into collaboration with USA on using the Brookhaven National Lab (BNL) relativistic heavy-ion collider facility (RHIC) in the Photon Electron New Ion Experiment (PHENIX) to

Box 4. Centre for Advanced Technology, Indore

A major centre for advanced technologies (CAT) – accelerator and laser related – has been set up at Indore by the DAE. Work on the first major accelerator facility consisting of a microtron, synchrotron and storage ring for production of high-energy electrons and associated synchrotron radiation is progressing very well at Indore. Major indigenous effort has gone into building this facility from design to fabrication (Figure B4). Sputter ion pumps of speeds 30, 70, 140 and 270 l/s have been developed at CAT for the accelerator and subsystems to achieve vacuum of the order of 10^{-10} mbar. All-metal Pirani and Penning gauges have also been developed to measure pressures from atmosphere to 10^{-9} mbar. Two vacuum furnaces for bulk degassing of SS components for accelerators have been developed at CAT. The furnace has a hot zone volume of 400 l which can attain a temperature of 700°C at 10^{-5} mbar.

The electron beam from 20 MeV microtron injector has been recently extracted. The synchrotron is fully assembled and all components well-aligned. All subsystems have been tested and this machine is expected to be operational in 1995. The transfer lines between the various machines have been assembled.

The electron beam from the synchrotron is being planned to be stored in Indus-1 in 1995. This will be a 450 MeV storage ring whose critical wavelength (defined as the median wavelength which divides the emitted synchrotron radiation energy into two equal parts) will be 61 Å. Useful flux can be obtained for wavelengths as short as 20 Å. The beam lines for the various experiments to be performed with Indus-1 are: (1) Photophysics (300–3000 Å), (2) Reflectivity measurements (40–1000 Å), (3) High-resolution spectroscopy, (4) Photoelectron spectroscopy, (5) L-edge spectroscopy, (6) EXAFS (100–1000 Å).

Indus-2 will have a critical wavelength of 4 Å from the bending magnets and a much lower wavelength of about 1 Å from the superconducting wiggler magnets. The major experimental programmes being planned for Indus-2 are: trace element analysis, lithography, small-angle scattering, diffraction, absorption-based spectroscopy, photoelectron spectroscopy and X-ray topography. The Indus-2 ring is expected to be commissioned in 1998.

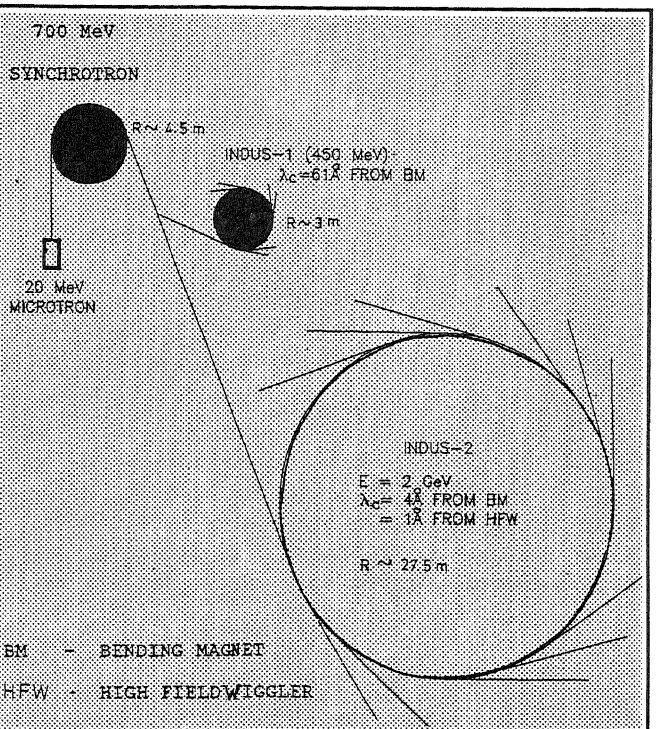


Figure B4. A schematic drawing of the synchrotron radiation source facility coming up at CAT, Indore.

search for the QGP. The BARC group will be involved in the design and the fabrication of the muon arm of PHENIX and also in the computer simulation related to this experiment.

The Indian scientists from TIFR and from Delhi and Chandigarh universities are also collaborating at the Fermi Lab, USA, on several frontier experiments and they are part of the teams which discovered recently the 'top quark'. In the future, India will also be collaborating in the currently planned highest-energy accelerator programme – the large hadron collider (LHC) – to be built in CERN.

Finally, the implication of all these laboratory experiments, in the context of the fundamental question of the origin of the universe a microsecond after the proposed 'big bang', when the universe must have consisted of quarks, gluons, leptons and photons, is in the postulation of the route to the synthesis of microcosmos and macrocosmos.

For several years, BARC has been collaborating with the Rutherford Appleton Laboratory in UK, which has a spallation neutron facility. The BARC group has built and installed a specialized neutron spectrometer in this centre – the so-called ΔT-window spectrometer – which uses the difference in temperature of two Be blocks to provide an ultrahigh energy resolution window for neutron inelastic scattering experiments.

Accelerator applications in industry and medicine

The charged-particle accelerators, primarily built for basic research in science, are being increasingly used for a variety of applications. Dedicated accelerators are also being built and operated for medical and industrial applications. The alpha particles from VECC have been used for characterization of (compound) semiconductors for oxygen. The helium-induced defects in refractory materials like Nb, Ta, W and V and in compound semiconductors like GaAs, GaP and InSb have been characterized. Information regarding migration of vacancies and the formation of vacancy clusters and helium bubbles has also been obtained from these studies. Some of these studies have direct implications for the mechanical and the metallurgical properties of refractory metals in a reactor environment. The helium jet recoil facility has been set up as a prelude to on-line isotope separation. Further, the alpha beams have been used for tribological investigations. Low-energy heavy-ion beams from Bhubaneswar have been used for materials science studies. Recently, using the proton backscattering tech-

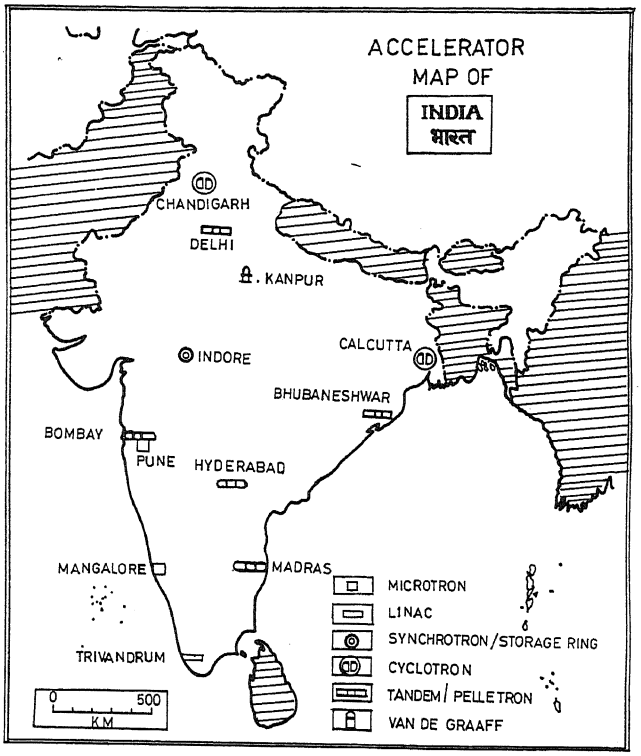


Figure 1. Accelerator map of India.

nique, the stoichiometry of thin VN films has been determined. Earlier, using the proton beam from Trombay Van de Graaff accelerator, the carbon build-up on the surface of a thorium target was estimated using the nuclear resonance technique. At the NSC, Delhi, a dedicated beam line has been developed for materials science studies, including high-energy heavy-ion irradiation to strengthen thin Cr films, characterization of superconducting films and bulk samples, and other ion-induced defects specifically involving high-energy heavy ions.

The use of particle accelerators in the study of problems in medicine and biology has been long recognized as a specific and powerful tool. The accelerators are being employed for both diagnosis and therapy. Certain specific radioisotopes like ^{67}Ga and ^{201}Tl required for accurate diagnosis of soft-tissue tumour and analysis of the functioning of the heart, respectively, can be produced only through nuclear reactions induced by charged particles from accelerators. Small quantities of ^{67}Ga are now routinely produced at VECC and used for diagnosis work in hospitals. Using the proton beam from VECC it is expected that the first batch of ^{201}Tl will be produced soon. A high-current captive cyclotron will be ideal to produce these isotopes at a reasonable cost. Trace element analysis of biological tissues, blood, etc., is another important area of investigation carried out using small-size accelerators.

The technique of radiation sterilization nowadays is used on a commercial scale in the manufacture of such products as disposable hypodermic syringes, blood transfusion kits, blood oxygenerators, Petri dishes, gloves, surgeon's needles and scalpels, besides some pharmaceutical products. Other areas include hygienization of sewage sludge and preservation of perishable food items. Electrons offer a substitute for gamma rays with the advantage of flexibility and inherent safety through on/off switching of the beam. A 10 MeV electron beam has been found to be suitable for this type of work. In India we are building a machine for this purpose. From economic considerations, a gamma ray source is favourable for industrial processing if source requirement does not exceed 1 million curies of ^{60}Co. The electrons complement the gamma rays from radioactive sources in many areas.

Radiation processing technology using electron accelerators has a pivotal role in an industrial environment. Low-energy electron accelerators having high beam intensity are being increasingly used in the radiation processing industry for a variety of applications.

At BARC, using the electron accelerator the following programmes are being pursued:

- Crosslinking of polythene 'O' rings for drum fittings – MOU with a Bombay firm for this application has been signed.
- Crosslinking of PVC/PE cables and wires for high performance and speciality applications.

There are plans to develop electron accelerators indigenously in BARC for this type of work.

For the radiation processing industry, plans are also under way at CAT to build a 750 keV, 20 kW electron accelerator mainly for the following items:

1. Radiation-processed crosslinked polyethylene cables and wires required by TV and computer manufacturers. Railways also import Rs 100 crores worth of these cables per annum.

2. Development of heat-shrinkable tubes. Telecommunication industry imports this item costing annually more than Rs 30 crores. CAT is also developing a 10 MeV, 15 kW electron Linac for industrial applications such as sterilization of medical products, irradiation of food products, etc. The dose rate available from such accelerators will be an order of magnitude higher than that from a typical ^{60}Co facility used for this purpose.

In order to demonstrate the economic viability of these various applications of accelerators in our country, CAT is thinking in terms of a medical and industrial accelerator centre which will be available to potential users for developing radiation-processing technology.

Use of radioactive sources like ^{60}Co for treatment of tumours is a well-known method. Recently, the electrons from low-energy electron accelerators are also being used for such treatment, particularly for superficial tumours. In India the radiation therapy based on accele-

rators is growing rapidly and at present there are about 20 odd electron accelerators operational in the country for this purpose. CAT has developed and operated a 20 MeV microtron as injector to a 700 MeV booster. Based on the experience gained on microtron, a 12 MeV microtron has been designed for applications like radiation physics studies and non-destructive testing and the same has been installed at Mangalore university. With addition of a Gantry, it is possible to make a radio therapy machine using the 12 MeV microtron. CAT is also at an advanced stage of building a 11 MeV cyclotron to produce short lived radio isotopes for positron emission tomography.

There exist also possibilities for use of accelerators in the nuclear fuel cycle, both for disposal of radioactive waste and for use as an 'energy amplifier' by introducing neutrons produced in spallation reactions into a subcritical assembly.

Summary

A possible strategy for accelerator-based research and applications can be summarized as follows:

- Research with existing facilities and utilization of the facilities for applied areas.
- Building new accelerator facilities for research and applications. Concentrating on reasonably large-size accelerators involving high technology but at moderate cost, which will enable work to be carried out in many frontier areas.
- International collaborations – physics research, participation in building accelerators and detector set-ups.

It is being realized that at least for research using big accelerators, cooperation amongst countries is very much needed not only for building expensive facilities but also for using the same. As we said before, no nation can any longer afford to go at such activities alone. However, our own national base has to be sufficiently strong and several national accelerator facilities are essential to provide this base. Our success in the international arena depends crucially on our expertise and competence built around our national facilities and development of manpower required to build and use the facilities both in India and abroad. A strong synthesis of the two programmes, national and international, should be the vision for our accelerator programme.

ACKNOWLEDGEMENTS. It is a pleasure to thank many of our colleagues from the DAE family and outside for providing valuable inputs to this report. In particular, we thank Dr D. D. Bhawalkar, Dr S. Gangadharan, Prof. S. N. Ganguli, Dr S. S. Kapoor, Mr S. S. Ramamurthi and Dr B. C. Sinha for valuable comments on the draft manuscript.

Recent fertility trends and prospects in India

K. Srinivasan

There is an increasing pace of fertility decline in large parts of the country in the recent years. Among the proximate determinants, the variables that have played a dominant role in fertility changes directly, in the Indian context are natural fertility and contraceptive use and indirectly, female literacy and infant mortality. For future, we can expect the TFR to be in the range of 2.9 to 3.0 by the year 2001 and 2.00 to 2.13 by the year 2011. The spurt in the female literacy rate will have a major impact on the future fertility levels.

THE most recent census[1] conducted in India in March 1991 placed her total population at 846.30 million, including an estimated figure of 7.72 million for the state of Jammu & Kashmir, where the census could not be conducted owing to political disturbances. Since 1951, when the first post-independence census was taken, the population of the country has increased by 485.22 million, more than twice the population that lived in the country in 1901, which was 238.40 million. During the decade 1981–1991 we have added to the country a whopping 162.97 million people, which is equal to the combined populations of France, Poland and the United Kingdom.

In terms of the rate of growth rather than numbers, during the decade 1981–91 the population has increased by 23.85%, or 2.12% annually, compared to 24.80%, or 2.20% annually, during the decade 1961–71. During the year 1991 the growth rate was 1.97%, almost the same as the average annual growth rate of 1.96% experienced 40 years earlier during the decade 1951–61. However, while the earlier growth rate was the result of a difference of an average crude birth of 42.4 and

K. Srinivasan is in the Population Foundation of India, B-28, Qutab Institutional Area, Tara Crescent, New Delhi 110 016, India.

death rate of 22.8 (birth and death rates measured for 1000 population per year) in the decade 1951–61, the 1991 growth rate was the result of a birth rate of 29.5 and a death rate of 9.8. Thus, during the 35-year period from 1956 to 1990 both the birth and death rates have declined by 13 points, keeping the growth rate almost unchanged. Unfortunately, this stagnation in the growth rate since 1961 has occurred, in spite of a strong awareness of the population problem at the national level since the country's first five-year plan in 1951, consistent antinatal policies and family planning programmes by the government to reduce the fertility levels in the population. An expenditure of approximately Rs 5550 crores has been spent on family planning since its inception until 1990.

This seems to be a paradoxical situation, in which the fertility declines have been tardy in a population when both at the people's level as well as at the governmental policy and planning level there has been a strong motivation to lower fertility levels and population growth rates. Many countries such as China, Thailand, Indonesia, Mexico and Brazil that launched their family planning programmes almost a decade later have achieved lower fertility levels than India. In 1991, China with a population size of 1155.5 million added to her numbers 16.52 million compared to 16.67 million added to India's. India has gained the unique distinction of being the country that adds the maximum number of people every year within her territory, surpassing China, which has a larger population. In order to understand the dynamics of this growth and prospects for the future, it is necessary to go into the components of population change, mortality, fertility and migration and the factors that have contributed to modifications in these components over time. Particularly relevant are the factors that influence the fertility levels in the Indian population; the component that holds the key to the demographic future of the country.

Fertility trends since 1970

Since 1970, fertility has been declining in most parts of the country as indicated by fertility measures. Table 1 presents data on two fertility measures, crude birth rate (CBR) and total fertility rate (TFR) and also on three important contributory factors to fertility change, viz. infant mortality rate (IMR), female literacy rate (FLR) and effective couple protection rate (ECPR or percentage of couples with wife in the age range 15–44 years effectively protected by a modern method of contraception from within the programme) in the country and in sixteen large states (that account for 95% of the country's population), for the years 1971, 1981 and 1991. The table also provides the CBR and IMR values for the latest available year 1993. This spans a period when the fertility transition has been occurring in the

Table 1. Measures of fertility, IMR, contraceptive use and FLR in India and the major states for 1971, 1981 and 1991

	CBR				TFR			IMR				ECPR			FLR		
	1971*	1981*	1991*	1993	1971**	1981*	1991*	1971*	1981*	1991*	1993	1971	1981	1991	1971	1981	1991
Andhra Pradesh	35.4	31.3	25.6	24.3	4.7	3.9	3.0	106	86	71	64	11.2	26.2	44.3	19.2	24.2	33.7
Assam	37.9	33.0	30.5	29.5	5.2	4.1	3.4	139	104	78	81	4.6	18.5	28.2	22.8	NA	43.7
Bihar	32.3	38.1	32.0	32.0	5.0	5.7	4.6	NA	118	72	70	5.0	12.3	26.0	10.7	16.5	23.1
Gujarat	40.4	34.8	28.4	28.0	5.6	4.4	3.2	144	113	69	58	11.4	32.5	57.8	30.3	38.5	48.5
Haryana	39.6	36.8	32.3	30.9	6.6	5.0	3.9	72	99	71	66	12.9	28.3	56.6	18.6	26.9	40.9
Himachal Pradesh	34.4	32.0	28.0	26.7	4.4	4.0	3.1	113	75	70	63	7.6	25.6	52.1	24.7	37.7	52.5
Karnataka	32.0	27.9	27.1	25.5	4.2	3.6	3.1	95	68	73	67	8.9	23.4	46.9	25.7	33.2	44.3
Kerala	31.3	26.2	18.5	17.4	4.1	2.9	1.8	58	36	17	13	14.6	29.7	55.6	65.4	75.7	86.9
Madhya Pradesh	39.2	37.7	35.9	34.9	5.9	5.2	4.6	135	139	111	106	9.2	21.5	40.3	13.7	19.0	28.4
Maharashtra	32.0	29.1	26.3	25.2	4.5	3.7	3.0	105	75	59	50	15.2	34.9	56.2	32.4	41.0	50.5
Orissa	35.8	32.7	28.9	27.2	4.6	4.2	3.3	127	137	120	110	13.8	27.1	41.0	17.0	25.1	34.4
Punjab	34.2	30.2	27.5	26.3	5.3	4.0	3.1	102	82	57	55	16.8	24.1	75.8	31.3	39.6	49.7
Rajasthan	41.1	37.9	34.5	35.1	6.2	5.4	4.5	123	103	84	82	5.0	13.8	29.0	10.5	14.0	20.8
Tamil Nadu	31.3	27.8	21.0	19.5	4.0	3.4	2.2	113	89	58	56	12.4	27.8	57.3	32.3	40.4	52.3
Uttar Pradesh	44.5	39.2	35.9	36.2	6.6	5.8	5.2	167	152	98	94	5.5	11.1	35.5	13.0	17.2	26.0
West Bengal	NA	32.4	26.7	25.7	NA	4.2	3.2	NA	91	66	58	NA	22.2	33.7	27.8	36.1	47.2
All India	37.2	33.8	29.6	28.7	5.2	4.5	3.7	129	106	80	74	10.1	22.8	44.1	22.9	29.8	39.4

Notes: (i) Figures indicate average rates over three consecutive years centred round the years 1971, 1981 and 1991. For example, CBR* for 1981 is the average of CBRs for 1980, 1981 and 1982. The rest of the figures are for the years 1971, 1981, 1991 and 1993.
(ii) TFR** for 1972 is based on SRS sample survey in 1972 and taken from ref. 6. The data are given for rural and urban areas and weighted by the 1971 rural–urban proportions in the census.
(iii) ECPR: Effective couple protection rate as on March 1971, 1981 and 1991.
(iv) FLR: Female literacy rate is percentage of literate among females aged 7 and above.
(v) 1993 figures are provisional estimates published by the Registrar General in September 1994.

Sources: (1) Various publications of the Sample Registration System by the Registrar General, including refs. 4 and 5.
(2) Ref. 3.

country. The fertility and mortality figures have been compiled from the official publications[2–6] of the Registrar General of India from the Sample Registration System, the female literacy rates from the 1971, 1981 and 1991 censuses and ECPR values from the Family Planning Year Books[7,8] put out by the Ministry of Health and Family Welfare. Since the Sample Registration System (SRS) is a dual record system based on data compiled on births and deaths from a sample of villages and urban blocks in the country (and though the sample sizes are fairly large enough to estimate fertility and mortality levels within 5% coefficient of variation), we have used three-year averages of vital rates, wherever available, centred around 1971, 1981 and 1991, especially for the study of long-term trends in these rates.

The CBR, defined as number of births per 1000 population per year, has declined from 37.2 in 1971 to 33.8 in 1981 and 28.7 by 1993. Thus, the annual decline in the CBR during 1981–93 was 0.425 compared to 0.340 during 1971–81, or 25% more than the earlier period. In 1993, among the states the CBR has ranged from a minimum of 17.4 in Kerala to a maximum of 36.2 in Uttar Pradesh. The CBR values for the rural and urban areas for 1993 for the country as a whole were 30.4 and 23.7, respectively. In Kerala the rural and urban CBR values were 17.5 and 17.3 respectively, while in UP they were 37.2 and 31.1 respectively.

It has been found that as fertility levels decline, the differences in fertility between the rural and urban areas as well as between the other segments of the population, such as by religion, caste or income categories, tend to narrow down and fertility becomes more homogeneous across the population. The more recently conducted National Family Health Survey (NFHS) that covered a representative sample of 88,562 households from most of the country has estimated the CBR for 1992–93 at 28.9, with 30.7 in rural areas and 28.9 in urban areas, which varied from a minimum of 17.2 in Goa and 19.6 in Kerala to a maximum of 36.0 in Uttar Pradesh, figures that are close to the estimates from the SRS, and provide independent validation to the SRS figures.

The changes in fertility or any variable can also be studied in relation to the progress made towards reaching the long-term goal set for different measures by an 'index of progress' (IP), by relating the actual decline in a fertility measure over a period of time to the required decline to reach the specified long-term goal in the measure. The recently appointed 'Expert Group

Table 2. Percentage progress towards long-term goals during 1971–81, 1981–91 and 1971–91 in CBR, TFR and ECPR

	Percentage changes in CBR (goal of 21)			Percentage changes in TFR (goal of 2.1)			Percentage changes in IMR (goal of 30)			Percentage changes in ECPR (goal of 70%)		
Long term goals	1971–81	1981–91	1971–91	1971–81	1981–91	1971–91	1971–81	1981–91	1971–91	1971–81	1981–91	1971–91
Andhra Pradesh	28.5	39.6	68.1	30.8	34.6	65.4	26.3	19.7	46.1	25.5	30.8	56.3
Assam	29.0	14.8	43.8	35.5	22.6	58.1	32.1	23.9	56.0	21.3	14.8	36.1
Bihar	–51.3	54.0	2.7	–24.1	37.9	13.8	NA	NA	NA	11.2	21.1	32.3
Gujarat	28.9	33.0	61.9	34.3	34.3	68.6	27.2	38.6	65.8	36.0	43.2	79.2
Haryana	15.1	24.2	39.2	35.6	24.4	60.0	–64.3	66.7	2.4	27.0	49.6	76.5
Himachal Pradesh	17.9	29.9	47.8	17.4	39.1	56.5	45.8	6.0	51.8	28.8	42.5	71.3
Karnataka	37.3	7.3	44.5	28.6	23.8	52.4	41.5	–7.7	33.8	23.7	38.5	62.2
Kerala	49.5	74.8	124.3	60.0	55.0	115.0	78.6	67.9	146.4	27.3	46.8	74.0
Madhya Pradesh	8.2	9.9	18.1	18.4	15.8	34.2	–3.8	26.7	22.9	20.2	30.9	51.2
Maharashtra	26.4	25.5	51.8	33.3	29.2	62.5	40.0	21.3	61.3	35.9	38.9	74.8
Orissa	20.9	25.7	46.6	16.0	36.0	52.0	–10.3	17.5	7.2	23.7	24.7	48.4
Punjab	30.3	20.5	50.8	40.6	28.1	68.8	27.8	34.7	62.5	13.7	97.2	110.9
Rajasthan	15.9	16.9	32.8	19.5	22.0	41.5	21.5	20.4	41.9	13.5	23.4	36.9
Tamil Nadu	34.0	66.0	110.0	31.6	63.2	94.7	28.9	37.3	66.3	26.7	51.2	78.0
Uttar Pradesh	22.6	14.0	36.6	17.8	13.3	31.1	10.9	39.4	50.4	8.7	37.8	46.5
West Bengal	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
All India	21.0	25.9	46.9	22.6	25.8	48.4	23.2	26.3	49.5	21.2	35.6	56.8

Notes: (i) The percentage progress made is measured as a ratio of the change in given period to the required change to reach the goal. For example the progress in CBR during 1971–81 is 100 * (B71 – B81)/(B71 – 21), where B71 and B81 are the CBRs for 1971 and 1981 and 21 is long-term goal for CBR.

(ii) The percentage progress made during two periods for the same area is additive; i.e. the progress made during 1971–91 is the sum of the progress made during 1971–81 and 1981–91.

(iii) *For TFR 1971–81 and 1971–91 the changes are from 1972 to 1981 and 1972 to 1991. The rest of the figures are for years 1971, 1981, 1991 and 1993.
**The ratio of change in 1981–91 to 1971–81 is not computed if there is an increase in CBR, TFR or IMR values during a period.

(iv) The long-term goals for all the states are assumed to be the same though they may be achieved in different years.

Sources: (1) Various publications of the Sample Registration System by the Registrar General, including ref. 5.
(2) Ref. 3.
Computed from figures given in Table 1.

on Population Policy' under the chairmanship of M. S. Swaminathan that submitted its report[9] in May 1994 has recommended the long-term demographic goals for TFR as 2.1 and IMR of 30 infant deaths per 1000 live births to be achieved by the year 2010. The Eighth-Plan document (1992–97) has recommended a long-term goal for a CBR of 21 to be reached by 2006–11, that implies an effective couple protection rate (ECPR) of 70%. The IP with regard to decline in the CBR value during 1971–91 towards the long-term goal of 21 can be measured as 100 × (CBR-71 minus CBR-91)/(CBR-71 minus 21).

Table 2 provides data on the IP values computed for major states for selected measures of fertility, mortality and contraceptive use in relation to their long-term goals. For the country as a whole the IP value for CBR during 1971–81 was 21% [100 * (37.2 – 33.8)/(37.2 – 21)], during 1981–91 was 28% and during 1981–91 was [100 * (33.8 – 29.5)/(37.2 – 21)]. Thus, during the whole period 1971–91, IP was slightly less than half (49%) towards the long-term goal of 21 set for CBR and the progress towards the goal during 1981–91 was 33% more than in the earlier decade. The IP values are studied only for the periods 1971–91 since we preferred to use three-year average values that are statistically more reliable around 1971, 1981 and 1991.

The decline in fertility during the period 1971–91 measured in terms of TFR, a more refined measure of fertility that denotes the average number of children born to a woman in her entire reproductive span assuming that she experiences the level of age-specific fertility rate obtained in a given year or period, is more pronounced. In 1991, TFR (average of 1991–93 values) was 3.7 births per woman in the country as a whole, ranging from a low of 1.8 in Kerala to a maximum of 5.2 in Uttar Pradesh. In the country as a whole, TFR declined from 5.2 in 1972 to 4.5 in 1981 to 3.7 in 1991. IP towards the long-term goal of 2.1 for TFR was found to be 49% during 1971–91, made up of 23% during 1972–81 and 26% during 1981–91. In this connection, it is worth mentioning that, as a starting point for the measure of TFR, we have used the estimate obtained from the Special Survey on Fertility and Mortality[10] conducted by the Registrar General in 1973, providing estimates for 1972, since those computed from the routine SRS, during the early seventies are considered to be underestimates[11]. Thus, while we have the change in TFR measured for a period of nine years during 1972–91, it is for a period of ten years during 1981–91.

IP in TFR has been more than 100% in Kerala, implying that the state had already achieved by 1991 the goal set for the country for 2010. The IP value was 96% in Tamil Nadu, which had almost reached the replacement level of fertility by 1991. It is, however, less than 50% in the states of Bihar, Madhya Pradesh, Rajasthan and Uttar Pradesh, states that are euphemistically referred to as 'Bimaru' states. The problem of fertility control is thus largely confined to the four large Hindi-speaking states in the north. Figures 1 and 2 provide a graphic presentation of the states arranged according to their progress or IP values on CBR and TFR during 1971–81 (1972–81 for TFR) and 1981–91. From Table 2 it can be seen that the pace of change during 1981–91 is more than during 1971–81 excepting in Assam, Karnataka, Punjab and UP. In other words, there is an increasing pace of fertility decline in large parts of the country in the recent years.

Factors contributing to fertility change

Earlier studies on factors associated with fertility levels and changes in India and other developing countries have identified the determinants at two levels, viz.

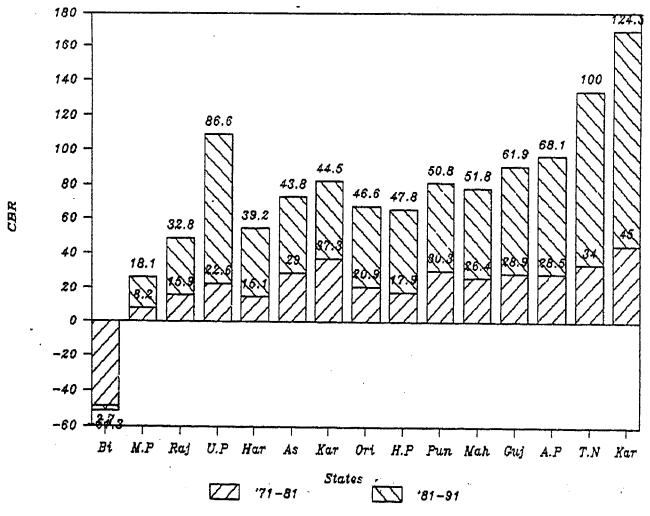


Figure 1. Percentage progress in CBR to goal 21 during 1971–81 and 1981–91.

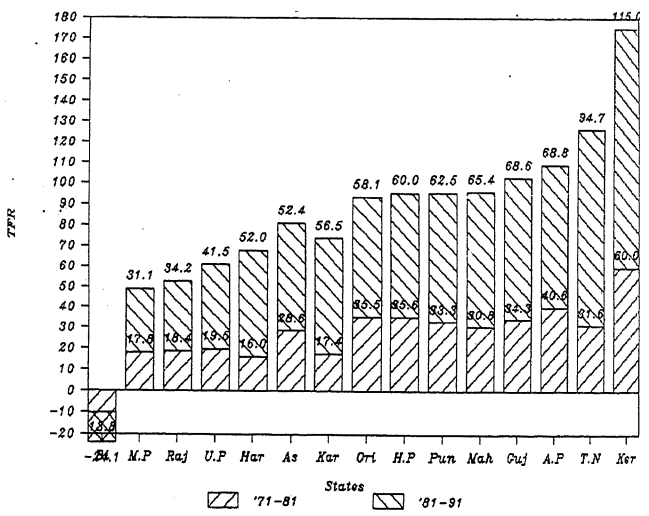


Figure 2. Percentage progress in TFR to goal 2.1 during 1971–81 and 1981–91.

proximate determinants and factors influencing proximate determinants. The proximate determinants of fertility are those intermediate variables through changes in any of which any modification in human fertility has to be brought about, viz. (a) *natural fertility* (NF), (b) *nuptiality* (NU), (c) *contraceptive use* (CU) and (d) *abortions* (AB) and three directly manipulable modernizing factors that are causally more distant from fertility as compared to the proximate variables, but that affect fertility powerfully through one or more of the proximate determinants, viz. (a) *family planning programmes* (FPP), (b) *female literacy* (FLR) and (c) *infant mortality* (IMR)[12–14]. Among the proximate determinants, the variables that have played a dominant role in fertility changes (increase or decrease) directly in the past in the Indian context are natural fertility and contraceptive use and indirectly, female literacy and infant mortality. We will briefly discuss their contributions one by one in the Indian context, and what they portend for the future of fertility in this country. To discuss the contributions of all the proximate determinants and their determinants will go beyond the scope of this article.

Proximate determinants of fertility

Natural fertility. Natural fertility is defined as the fertility of a married couple in the absence of any deliberate control to increase the spacing between children or to limit the family size. In India, the natural fertility has been historically quite low, probably the lowest in the world. It was only about six births per woman, or just 60% of the number observed in European populations in the eighteenth and nineteenth centuries, when they did not practice any contraception. Fertility levels have been kept in check in India by various social and cultural practices and to some extent are attributable to the high prevalence of febrile diseases such as tuberculosis, malaria and malnutrition that reduced the biological fecundity and in some cases caused infertility in the couples. Despite the tradition of early and universal marriage, various religious and cultural restrictions on sex, prohibitions against widow remarriage, prolonged breast feeding contributed to low natural fertility. With modernization gaining rapid momentum after independence, these traditional and cultural practices that kept the natural fertility of Indian population at low levels for centuries have weakened considerably. Further, the increases in life expectancy and improvements in nutritional and health status attributable to green revolution and control of many communicable diseases, the duration of married life that a woman spends in the reproductive period has increased rapidly and the natural fertility levels have increased steadily to approximately 7 children in the seventies to 9 in the eighties[14–16]. From a detailed analysis of data on age-specific marital fertility rates compiled from two large-scale sample surveys conducted in 1959 and 1972 in 11 major states in India, Srinivasan and Jejeebhoy[17] observed:

'The results consistently indicate an increase in the total marital fertility rates in eight of the eleven states (the exceptions being Kerala, Orissa and Punjab) between 1959 and 1972. The most striking increases in marital fertility have been observed in those states where there has been practically no prevalence of contraception. The increase is 28% in Uttar Pradesh, 26% in Rajasthan, 13% in Madhya Pradesh and all these states were practically under natural fertility conditions . . .' (p. 103).

This tempo of rise in natural fertility initiated in the early sixties continued into the seventies and in the state of Bihar the rising trend is observed even in the eighties. Pathak and Ram[16], using indirect techniques, estimate the level of natural fertility to be slightly more than 9 children in the eighties. Thus, it has to be recognized that whatever decline in the birth rate that has occurred in India has taken place in the context of an almost steady rise in the natural fertility levels of the population to about one-and-a-half times the initial level. The plateauing of this rise seems to have occurred in the eighties, and there may not be any further fertility-increasing effects of modernization. It is to be noted that in the absence of the vigorous family planning programme that we had in the past few decades, the observed fertility levels in the Indian population would have risen considerably.

Contraceptive use. India can take legitimate pride in the fact that it is the first country in the world, developed or developing, to realize at the governmental level the extent of hurdles to development posed because of rapid population growth and launch an official programme of family planning as early as 1951 to control fertility as a part of the first five-year plan (1951–56). The governmental emphasis on the programme has accelerated over the years with plan expenditure on the programme rising sharply from a meagre Rs 14.5 lakhs in the first plan to 2304.5 crores in the seventh plan period, 1985–90. The current level of annual expenditure on the programme incurred by the states and the centre (plan and non-plan), mostly by the centre, is about 1000 crores per year. In 1991, there were about 131,000 subcentres, 22,060 Primary Health Centres, 1930 Community Health Centres, 1515 Urban Family Welfare Centres and about 554 district/state-level hospitals providing family planning services across the length and breadth of the country.

The impact of such a massive expansion of activities in the field of family planning over the years is seen from the data on the increasing prevalence of contraceptive use in the country (given in Table 1). The effective couple protection rate (ECPR) or the percentage of married women in the reproductive ages (15–44)

protected effectively from pregnancy by a modern method of contraception has increased from almost zero level in the late fifties, to 10.1% in 1971, 22.0% in 1981 and 44.1% in 1991. In actual numbers, 64.07 million couples are effectively protected by contraception under the programme. Out of this 44.1% of couples using contraception effectively in 1991, 30.3% (or 68% of those protected) have been protected by sterilization (of the husband or wife), 6.7% (or 15% of those protected) by IUD and 7.2% (17% of those protected) by other methods that include oral pills, condoms, injectibles and others. Thus, sterilization, vasectomy until 1976 and tubectomy thereafter, has been the dominant method promoted by the programme from its very inception and accounts for the bulk of the protection in the country.

In 1991 among the states the ECPR values ranged from a low of 26.0% in Bihar to 75.8% in Punjab. The IP values in ECPR furnished in Table 2 reveal that while towards the long-term goal of achieving 70% in the ECPR values, only 21.2% progress was made during 1971–81; during 1981–91, 35.6% was made and the ratio of progress between the two decades is thus 1.7. This ratio has ranged from 7.1 in Punjab (where the progress during the 1971–81 decade was poor) to 0.7 in Assam. In all the states other than Assam, the IP values in 1981–91 are higher than in the previous decade, indicating a rapid acceleration in the use of contraception in recent years. A considerable degree of caution has to be exercised in projecting this trend to future years because of the various assumptions made by the Department of Family Welfare in the computation of ECPR values. However, even if we assume that the quantum of increase in the ECPR values experienced during 1981–91, i.e. 21.8%, will be maintained during 1991–2001, it will lead to an ECPR level of 65.1% in the year 2001 for the country as a whole, a level quite close to what is required to realize the replacement level of fertility. It has to be realized that further increase in ECPR values has to come mostly from the hitherto poorly performing states of Assam, Bihar, Rajasthan, Orissa, Madhya Pradesh and Uttar Pradesh, that have performed below the national average in 1991 and that have about half the population of the country.

An additional positive feature of changes occurring in the country during the past two decades is the reduction in the gender bias of the population. The people of this country have for centuries placed undue emphasis on the male child and this gender bias has been institutionalized in a number of ways through religion and other social institutions. The female child has been given a secondary position in the family and society in all aspects, including education, property ownership, health care and decision making. This bias seems to be breaking down in recent times. The data collected in a large-scale survey[18] conducted during 1988–89 by the Operations Research Group (ORG), Baroda, to study the levels, trends and determinants of contraceptive use in the country revealed that among couples in the reproductive ages (wife in the age range 15–44), among those with two surviving children, both daughters, 35.6% were using a modern method of contraception and 15.6% have undergone sterilization. Even among those with no surviving children 12.0% were using a modern method of contraception. Figure 1 provides the contraceptive prevalence rate by numbers of surviving children and surviving sons. While there is an increased use of contraception among couples with more sons, the practice among those without any son has also risen substantially.

One of the major reasons for the reduced impact of the family planning programme on the fertility levels in India is the undue emphasis placed by the programme on the method of sterilization, vasectomy for males until 1976 and tubectomy for females thereafter. Since this is largely an irreversible and permanent method, couples prefer to adopt it after they have reached the desired family size and at older ages and higher parities. Couples who wanted to space their children and keep the option of their fertility open did not take to these methods and invariably landed up with a larger family size than what they would have liked to have if good temporary, reversible methods were made available to them from the early years of the programme. The sterilization programme has been like 'locking the stables after the horses have been bolted away'. Further, with the rise in natural fertility pushing up the fertility levels of the younger married couples (because of relaxation of traditional checks on fertility and control of febrile diseases and better nutrition) only the tail end of the fertility curve was affected by the programme while at the peak ages of fertility, viz. 20–29 years, the marital fertility was rising during the sixties and seventies. Since the rising phase of natural fertility had largely been completed by the late eighties (possibly with the exception of Bihar) and increasing emphasis is being placed on the spacing methods in recent years, the fertility impact of the programme is likely to increase in the coming years with more and more younger women coming to use modern methods of contraception.

Another indication of major changes taking place in the population is on the extent to which couples have reported not desiring any more children at different stages of their family building in the early eighties and nineties. In a large-scale sample survey conducted by the ORG, Baroda, in 1980–81, the percentage of eligible couples who did not desire any additional children in the country as a whole was 1.0, 12.7, 43.3, 68.7 and 83.7, among married women with 0, 1, 2, 3 and 4 or more living children respectively. The corresponding percentages in the 1992–93 National Family Health

Survey[19] are 2.3, 19.1, 61.8, 78.0 and 83.6, respectively. These figures include a sizeable proportion who have already resorted to sterilization to avoid further child bearing. Particularly impressive is the rise in the proportion of women with two living children not wanting any more children, which is 62%, half of whom had already undergone sterilization.

Nonproximate developmental factors in fertility change

Declines in infant mortality. Detailed studies of the demographic transition in the developed and developing countries have revealed a strong link between declines in the mortality levels of a population, especially in the infant and child mortality and fertility levels. The well-known 'theory of demographic transition' is based on the empirical evidence from a number of countries that have successfully achieved the transition from a high-fertility–high-mortality to low-fertility–low-mortality regime. In all these cases, mortality declines always preceded fertility declines with a time lag and the pace of mortality decline determined the pace and timing of fertility decline even in the absence of any organized contraceptive services. Family planning programmes assist and expedite the pace of fertility decline once they have commenced as a part of the multiphasic social response to earlier mortality declines[20].

Life expectancy in India has improved substantially from the third decade of this century from an estimated 20 years in 1921 to 60 years in 1991. During the four decades 1941–51 to 1981–91, life expectancy rose by 25 years with about 0.6 years added to an Indian's life each year. Credit for this considerable achievement goes to various public health measures undertaken by the central and state governments. Life expectancy varies considerably among the states, with 71.1 years in Kerala to 49.6 years in Uttar Pradesh for the quinquennium 1986–91. The crude death rate has declined from 47.2 fatalities per thousand population during 1911–21 to 9.3 in 1993. The decline in the death rate seems to have begun after 1921, and since 1951 has been quite rapid. Infant mortality rate (IMR; deaths of children below age one year per 1000 live births), was estimated at 222 for the decade 1901–11, declined to 129 for 1961–71 and to 74 by 1993. There are sizeable interstate variations in the IMR values. In 1993, IMR ranged from 110 in Orissa to 13 in Kerala. There are also rural–urban differentials in the IMR values, 82 in the rural areas compared to 45 in the urban areas in 1993. The differentials can be largely attributed to the prevalence of better health and medical care facilities in states such as Kerala, Maharashtra, Tamil Nadu and West Bengal and to the relative concentration of these facilities in the urban areas in most states. It has been observed that as the infant mortality levels decline, rural–urban differentials tend to narrow down. For example, in Kerala, in 1993 the rural IMR was 15 compared to 8 in the urban areas while in Orissa the rural rate was 115 compared to 69 in urban areas. The reduced differentials in mortality seem to be an indicator of the extent of dispersion of basic health services across the country. The declines are attributable to the control of various infectious and communicable diseases that had been taking heavy tolls among children. Concomitantly, there is a tendency for mortality rates among the adults and the elderly to rise because of increases in noncommunicable diseases, including cardiovascular problems, cancer and accidents. The shift from the age pattern of deaths from younger to older ages and predominantly infectious to noninfectious diseases has been termed as the 'epidemiological transition' and many states in India are almost completing such a transition. The basic requirement for a sustained fertility transition, viz. decline in the mortality levels, especially the infant and child mortality levels, is present in the Indian population with varying degrees of success in different states. The states that have achieved faster declines in mortality levels such as Kerala have achieved lower fertility levels well ahead of the other states.

Analysing the progress made in the reduction of IMR values towards the long-term goal of 30 infant deaths per 1000 live births, through the index of IP given in Table 2, it is found that in the country as a whole the IP during 1971–91 is about 50%, with 23% during 1971–81 and 26% during 1981–91. There is thus a 13% acceleration in the pace of decline in IMR values in recent years. Figure 3 gives a graphic representation of IP values in different states during 1971–81 and 1971–91. The minimum decline has been recorded in the state of Orissa and more than 50% IP values are found in

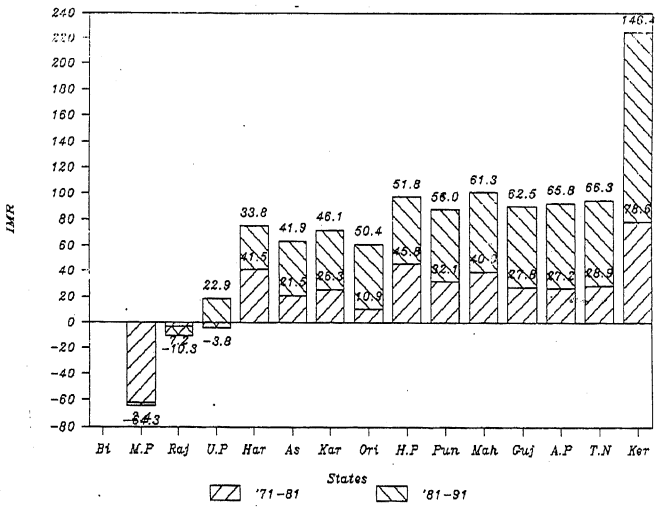


Figure 3. Percentage progress of IMR to goal 30 during 1971–81 and 1981–91.

the states of Assam, Himachal Pradesh, Kerala, Maharashtra, Punjab, Tamil Nadu and Uttar Pradesh. It is found that among the larger states only the state of Kerala has already achieved both the fertility and infant mortality goals set for 2010, with a TFR of 1.8 CBR of 18.3 and IMR of 16 by 1991. While Tamil Nadu has almost reached the fertility goals with TFR of 2.2 and CBR of 20.8 by 1991, its infant mortality rate of 56 in 1993 is still distant from the goal of 30. However, Tamil Nadu and Punjab have achieved the IMR goals set by the earlier National Health Policy of 1983, wherein the goal was to reach a level of 60 or less by the year 2000. Thus, the picture that emerges from an analysis of the changes in the IMR values between 1971 and 1991 is that the tempo of change (decline) in the IMR values has accelerated in the recent decade 1981–91 compared to the earlier decade and this is bound to have an effect in accelerating the pace of decline in fertility in the coming years.

Female literacy. One of the major determinants of demographic transition leading to declines in infant and child mortality, overall mortality, increase in age at marriage, rise in use of modern methods of contraception, declines in the fertility levels of the population and development *per se* is education of women. Societies that have achieved universal literacy of females and higher educational attainments for women have all recorded more rapid demographic transition. The fertility transition in China, Sri Lanka and Kerala state in India all bear ample testimony to the above conclusion. This is demonstrated strongly by data in Table 1, wherein a strong negative association between female literacy rate and fertility levels is observed in all the three census years (e.g. -0.85 in 1991).

Literacy rate for females aged 7 and above increased from 22.9% in 1971 to 29.8% in 1981 to 39.4% in 1991. In 1991 it ranged from a high of 86.9% in Kerala to a low of 20.8% in Rajasthan. While the literacy rate both among the males and females has been rising over the decades, the absolute number of illiterate females in the population aged 7 and over has been increasing steadily, to 126.7 million in 1991, with an increase of 5.7 million during the decade 1981–91. Almost 80% of the illiterate female population of the country (about 96 million) are concentrated in the four large Hindi-speaking states of Bihar, Madhya Pradesh, Rajasthan and Uttar Pradesh.

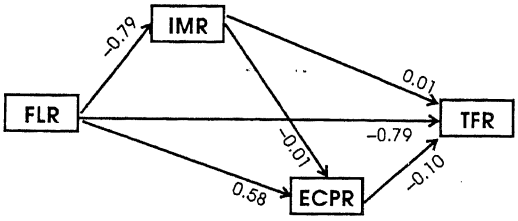


Figure 4. Path model.

A multivariate model

The combined and separate effects of three of the four contributory factors discussed above, viz. FLR, IMR, ECPR on the TFR can be assessed by a simple path model. We are not in a position to take account of the factor of natural fertility since we do not have direct data on the same and any measure of natural fertility has to be derived only indirectly from the observed fertility measures and the extent of the use of contraception.

In the path model (Figure 4) each variable is assumed to be influenced only by others posited to the left of it and not by any on its right. In the above model the TFR is assumed to be influenced by all the three factors – FLR, IMR and ECPR; ECPR is assumed to be influenced by IMR and FLR, and IMR by FLR, in addition to the other factors not explicitly entering into the model but affecting the variability of the factors. The zero-order correlation coefficients among the four variables for the years 1981 and 1991 are furnished in Table 3, where the figures in the upper triangle of the matrix provide the coefficients for 1981 and the lower triangle for 1991. Using the data for 1991 given in Table 1 the following regression equations emerge by applying the simple least-square method of estimation. All the correlation coefficients are statistically significant and reveal the strong interrelationships that exist among them.

Using the above path model and standardized variables, the regression equations among the variables are as follows. In this case the regression coefficients are the path coefficients, as per the above causal model.

$$\text{TFR} = -0.79 * \text{FLR} + 0.01 * \text{IMR} - 0.10 * \text{ECPR}, \quad (1)$$
$$n = 16 \text{ and } R\text{–squared} = 0.74,$$

$$\text{ECPR} = 0.58 * \text{FLR} - 0.01 * \text{IMR}, \quad (2)$$
$$n = 16 \text{ and } R\text{–squared} = 0.35,$$

$$\text{IMR} = -0.79 * \text{FLR}, \quad (3)$$
$$n = 16 \text{ and } R\text{–squared} = 0.63.$$

Table 3. Correlation matrix of selected variables in 1991

TFR	1.00000			
IMR	0.68190	1.00000		
ECPR	–0.56962	0.47290	1.00000	
FLR	–0.85386	0.79412	0.58982	1.00000

Critical value (1-tail, 0.05) = ±0.42706
(2-tail, 0.05) = ±0.49580

$n = 16$

The above analysis reveals that 74% of the variability in TFR is explained by the three factors considered but among them only FLR has a significant direct impact on fertility with a very high magnitude path coefficient of -0.79. Similarly, FLR has a statistically significant impact on the IMR as well as on the ECPR value. Thus, in the present context FLR seems to have a major effect, directly on TFR through its effect on age at marriage, use of induced abortions, etc., and indirectly on marital fertility rate through better acceptance of contraception from within the programme.

Interestingly, repeating the above analysis with the data for 1981 and using the standardized variables, the following equations emerge. Since the regression coefficients are based on standardized variables and they are the path coefficients, they permit comparison of the effects of variables over time and among the different variables.

$$\text{TFR} = -0.28 * \text{FLR} + 0.39 * \text{IMR} - 0.38 * \text{ECPR}, \quad (4)$$
$$n = 15 \text{ and } R\text{–squared} = 0.82,$$

$$\text{ECPR} = 0.63 * \text{FLR} - 0.01 * \text{IMR}, \quad (5)$$
$$n = 15 \text{ and } R\text{–squared} = 0.41,$$

$$\text{IMR} = -0.78 * \text{FLR}, \quad (6)$$
$$n = 15 \text{ and } R\text{–squared} = 0.78.$$

Comparison of the results for 1991 and 1981 reveals that the effect of FLR on IMR has remained stable and significant over time with over 75% of variance in IMR value explained by FLR. The effect of FLR on ECPR has declined over time. With regard to their direct impact on TFR values, while IMR and ECPR were having a significant effect and FLR not a significant effect in 1981, the effects are reversed in 1991 with only FLR having significant effect on fertility in 1991. This suggests that in the coming years more than the family planning programme, which offers the family planning services, education of women may play a more crucial role in shaping the fertility values in the country. The increased emphasis placed by the state and central governments on the literacy programme through the national literacy campaigns with emphasis on female literacy will have its direct and indirect impact of the fertility levels in the coming years.

Future prospects for fertility

As with any prediction exercises of future human behaviour, prediction of future fertility levels in India is a highly probabilistic endeavour, with the level of uncertainty increasing with the time span over which the predictions are made and the smaller level of population aggregation at which it is made. Predictions over a longer time span and for smaller geographic or population size zones are more loaded with uncertainty than predictions over a shorter time span and for the country as a whole.

Let us undertake this venture at two levels, one a simple and the other a more complex modelling. For example, TFR was 3.7 in 1991, 4.5 in 1981 and 5.2 in 1972.

During 1972–81 it declined by 0.7 points (5.2 to 4.7) while during 1981–91 by 0.8 points (4.5 to 3.7), implying that the reduction in TFR over the 20-year period was almost at the rate of 0.8 points per decade. If this trend of decline continues into the future, i.e. at 0.08 point reduction per year, then the TFR will be 2.9 by 2001 ($3.7 - 0.8$). Similarly, with regard to CBR, if we assume that the pace of decline during 1981–91 will continue into the future, the expected value in 2001 will be 25.5 ($29.6 - 4.2$). These are simplistic extrapolations of earlier trends.

We can make the projections of fertility a bit more complex and indirect by using regression of TFR on the important contributory factors discussed above, viz. IMR, FLR and ECPR and using the projected values of the predictors. The regression equation using the actual values (rather than standardized values as needed for the path model), based on the 1991 data is:

$$\text{TFR} = 5.6324 + 0.0004 * \text{IMR} - 0.0066 * \text{ECPR} - 0.0446 * \text{FLR} \quad (7)$$
(multiple $R = 0.8578$).

The projected values of the three predictors by the year 2001 can be obtained as follows under as realistic and conservative an approach as possible.

1. *Infant mortality rate.* For the country as a whole IMR has declined from 129 in 1971 to 106 in 1981 and to 80 in 1991. The governments at the state and central level have, in recent years, intensified their child immunization programmes throughout the country in collaboration with and assistance from the UNICEF under the Child Survival and Safe Motherhood Programmes (CSSM). There has also been tremendous media blitz on the need for immunization, and child care and nutrition, and we can expect, on a conservative assumption, that in the next ten years, during 1991–2001, the quantum of decline in the IMR values will be of the same magnitude as in the decade 1981–91, i.e. a decline by 26 points. Under this assumption we can project the IMR value in 2001 at 54.

2. *Effective couple protection rate.* For the country as a whole the ECPR values have increased during the decade 1981–91 by 21.3 points from 22.8 to 44.1%. With the increased emphasis on spacing methods and concern for improvement of the quality of services, in recent years, we can expect this tempo to be maintained in the period 1991–2001 and we can expect ECPR to increase to 65% ($44 + 21$) by the year 2001. This assumption seems to be a bit optimistic and hence we will

make another assumption, on the more conservative side, 55%, in which we assume that the increase in the ECPR value is only 11%, slightly less than what was achieved in the decade 1971–81, which was 12.7%. We will thus have two assumptions of 55% and 65% for the ECPR values by the year 2001.

3. *Female literacy rate.* FLR is going to be the cutting edge of demographic change, especially fertility transition, in this country in the years to come. It has increased from 22.9% in 1971 to 29.8% in 1981, to 39.4% in 1991, or by a factor of 1.30 and 1.32 in the two decades. In recent years there has been considerable emphasis on the literacy of women and the girl child through the national literacy campaigns, mid-day meal schemes for school children, and various mass media programmes. I expect such a national movement to increase FLR exponentially by the same factor between 1981 and 1991. In other words, I expect the literacy rate among females aged 7 and over by the year 2001 to be FLR-91 * (FLR-91/FLR-81) or 52.0.

Under these assumptions of future changes in the predictors of fertility, i.e. IMR of 54, ECPR of 65 or 55 and FLR of 52 by 2001, using equation (7) we get TFR in 2001 as 2.91 or 2.97, with ECPR at 65 or 55, respectively. With the same reasoning as above, if we extrapolate the predictor values to 2001, IMR will be 28, ECPR will be 86 or 66 and FLR will be 69 by that year. Using these values for the predictors in the regression equation we get a TFR for 2011 as 2.00 or 2.13 for the two ECPR alternatives. The long-term goal of TFR of 2.1 to be reached by the year 2011 will definitely be reached under these assumptions. Thus, we can expect TFR to be in the range 2.9–3.0 by the year 2001 and 2.00–2.13 by the year 2011. If there is going to be a spurt in FLR in the coming years (which is not unlikely), it will have a major impact on the future fertility levels. It has to be realized that a 20% point increase in the FLR value will reduce TFR by one child.

I have not discussed in this exercise the various difficulties that are going to be faced in the programmes for reduction of IMR, increase in the use of contraception within the programme (ECPR) or raising the FLR, and these are by no means to be underestimated. The need to intensify our efforts on immunization and child nutrition programmes, improvements in the quality of services offered in the maternal, child health and family planning programmes and achievement of universal literacy for all with emphasis on the schooling of the female child are to be emphasized at the risk of repetition. However, I wish to point out that fertility has started its downward trend in this country, that the pace of decline is accelerating in the past decade and that the changes in three important determinants of fertility, IMR, ECPR and FLR, are also changing fast enough to have synergistic dampening effect on fertility and that the goal of replacement level of fertility or TFR of 2.1 will definitely be reached by the year 2010 as targeted in the recent population policy document, if not earlier.

1. Census of India 1991, Paper 1 of 1991 – *Provisional Population Totals*, Registrar General & Census Commissioner, New Delhi, 1991.
2. Registrar General of India, Report on Levels, Trends and Differentials in Fertility, Vital Statistics Division, Ministry of Home Affairs, New Delhi, 1979.
3. Registrar General of India, Sample Registration System, Vital Statistics Division, Ministry of Home Affairs, New Delhi, 1990.
4. Registrar General of India, Sample Registration Bulletin, Ministry of Home Affairs, New Delhi, July 1994, vol. XXVIII, No. 2.
5. Registrar General of India, Fertility and Mortality Indicators 1992. Sample Registration System, Vital Statistics Division, Ministry of Home Affairs, New Delhi, 1994.
6. Registrar General of India, Fertility Differentials in India 1971, New Delhi, 1976, pp. 45, 48.
7. Department of Family Welfare, Family Welfare Programme in India, Year Book 1989–90, Ministry of Health and Family Welfare, Govt. of India, New Delhi, 1991.
8. Department of Family Welfare, Family Welfare Programme in India, Year Book 1990–91, Ministry of Health and Family Welfare, Govt. of India, New Delhi, 1991.
9. Expert Group on Population Policy, Draft National Population Policy, submitted to the Ministry of Health & Family Welfare, Government of India, New Delhi, May 1994.
10. Registrar General of India, Special Survey on Fertility and Mortality,
11. Bhat, M., 1994 (personal communication).
12. Bongaarts, J., *Popul. Dev. Rev.*, 1978, **4**, 105–132.
13. National Academy of Sciences, *Determinants of Fertility in Developing Countries* (eds Bulatao, R. and Lee, R.), Academic Press, New York, 1984.
14. Srinivasan, K., *Regulating Reproduction in India's Population: Efforts, Results and Recommendations*, Sage, New Delhi, 1995.
15. Srinivasan, K., Shireen, J. J., Richard, A. E. and Eiuleen, M. C., *Popul. Dev. Rev.*, 1984, **10**, 273–296.
16. Pathak, K. B. and Ram, F., in *Dynamics of Population and Family Welfare* (eds Srinivasan, K. and Pathak, K. B.), Himalaya Publishing House, Bombay, 1992, pp. 279–293.
17. Srinivasan, K. and Shireen, J., in *Dynamics of Population and Family Welfare* (eds Srinivasan, K. and Mukerji, S.), Himalaya Publishing House, Bombay, 1981, pp. 91–117.
18. Operation Research Group (ORG), Family Planning Practices in India: Third All-India Survey, Baroda, 1991.
19. International Institute for Population Sciences, National Family Health Survey 1992–93: India – Introductory Report, Bombay, October 1994.
20. Kingsley, D., *Popul. Index*, 1963, **29**, 345–366.

Adaptive features of the predatory ladybird beetles in uncertain and competitive habitats

B. K. Agarwala and S. Bhattacharya

Department of Life Science, Tripura University, Agartala 799 004, India

The predatory species of ladybird beetles forage in habitats where aphid colonies show a patchy distribution in space and time. In an uncertain environment the numerical response of the females to aphid density suffers from unstable food source and sharing of the habitat by different ladybird species. As a result, the reproductive effort of a predatory ladybird species is threatened by the intraspecific and interspecific competitions. It is imperative that ladybird species should synchronize their reproduction and development of progeny with phenology of prey populations in order to overcome the risks of their habitats. This paper brings together the results of several studies to present an account of the adaptive mechanisms evolved by the predatory species of ladybirds to cope with the risks inherent in their life cycle.

MANY species of ladybird beetles (Insecta, Coleoptera, Coccinellidae) are predators of aphids, and other insects and mites. Some of these ladybird beetles are regarded as important natural control agents because of their ability to feed on large quantities of aphids[1]. Several species of aphids are a menace to agriculture and horticulture[2–4]. As part of integrated pest management strategy, increasing efforts are being made to understand the ecology of predatory ladybird beetles in the development of successful biological control of pest aphids[5,6].

Field studies suggest that larvae and adults of most ladybird species search plant surfaces at random and concentrate their efforts on the restricted area where a prey is captured, i.e. they show a positive prey-taxis[7,8]. As a consequence of this foraging behavior, ladybird beetles manage to exploit aphid colonies that have a patchy distribution and a short life span[4]. This tendency results in an aggregation of either conspecific larvae and adults or other species of ladybirds that share the same habitat and, therefore, are likely to interact with one another[1,9,10]. In such a competitive environment, it is possible that the reproductive output of individual species is threatened by one or more of the following risks:

1. temporary nature of food supply – for instance, sudden collapse of aphid colonies before larval development is completed;
2. cannibalism of eggs and larvae; and
3. interspecific predation.

How predatory ladybirds overcome these inherent risks is the question this paper seeks to answer by reviewing the adaptive features of their reproductive behaviour, interspecific predations and cannibalism.

Reproductive behaviour

The reproductive behaviour of aphidophagous ladybird beetles is guided by the following facts: (i) the aphid populations persist for a relatively short period of time[4]; (ii) the developmental time of ladybirds and the life span of aphid colonies are approximately equivalent[11]. and (iii) there is necessity of avoiding any competition between larvae for the limited food resource. Further, it has been shown that the newly born larvae need a higher aphid density to survive than the older ones, and that feeding success of these young larvae is negatively correlated to prey size[12]. In these conditions it is conceivable that the life cycle of ladybirds should synchronize with the development of prey populations.

Field observations based on survey of different habitats occupied by *Adalia bipunctata* L. show that eggs are generally laid in aphid colonies early in their development[12,13] and oviposition is limited to a short period of the early phase of population growth[14]. Similar results have been obtained for another ladybird species *Coccinella transversalis* (Fabricius) in bean crops infested by *Aphis craccivora* Koch[1] (Figures 1 and 2).

Laboratory observations further suggest that oviposition is strongly inhibited by the presence of conspecific larvae or other females[14] (Figure 3). Such females not only produced less number of eggs but also took more time to lay eggs than control females. Interestingly, however, ovipositing females did not respond adversely to the presence of their own pupae and eggs. The possible cause of such an interference could be some form of signal, as reported in parasitoids[5], to the ovipositing female, who gets the indication that the size of aphid colony is not suitable for egg laying.

A recent study[6] suggests that an adult ladybird female lays an optimum number of eggs in an aphid colony. Since an aphid colony is a finite and temporary food source, the time of its existence is just enough for the larvae of ladybird to complete their development but not reproduce in that colony (Figure 4)[16]. This implies

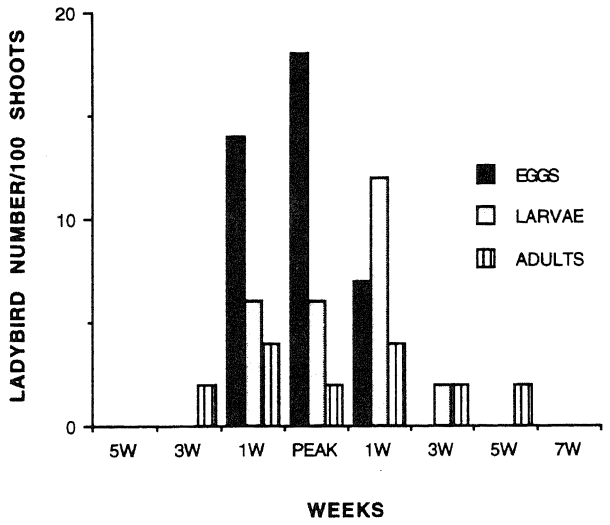


Figure 1. Distribution of ladybird eggs, larvae and adults of *Coccinella transversalis* in bean fields in time relative to peak abundance of aphid *Aphis craccivora*.

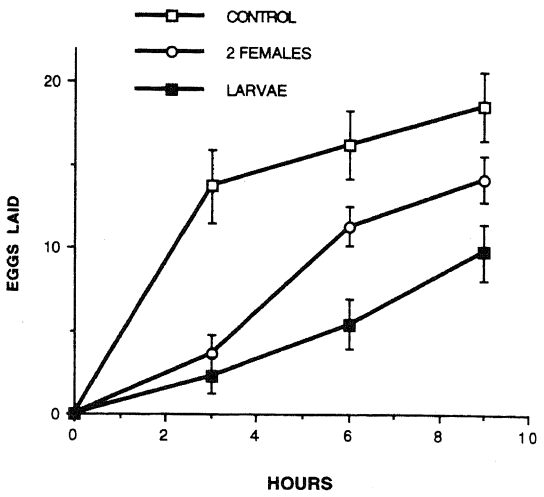


Figure 3. Average number of eggs laid by the females of ladybird *Adalia bipunctata* during 3, 6 and 9 h in 5 cm Petri dishes, in which they were kept singly, with another female or with three fourth-instar larvae (30 replicates for each treatment) (after ref. 14).

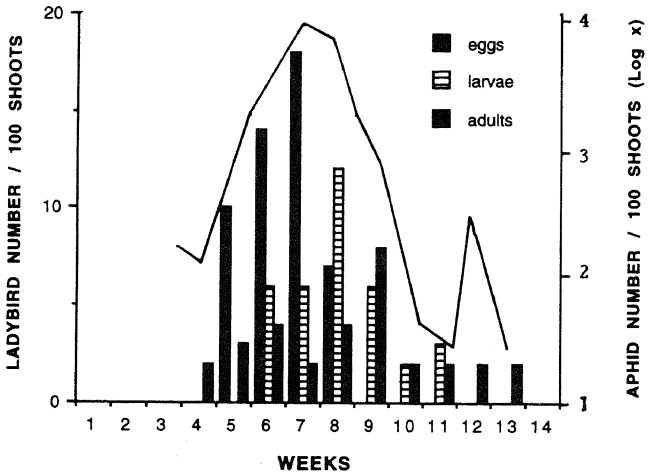


Figure 2. Number of aphid *A. craccivora* (—) and ladybird *C. transversalis* (histogram)/100 shoots recorded in bean field.

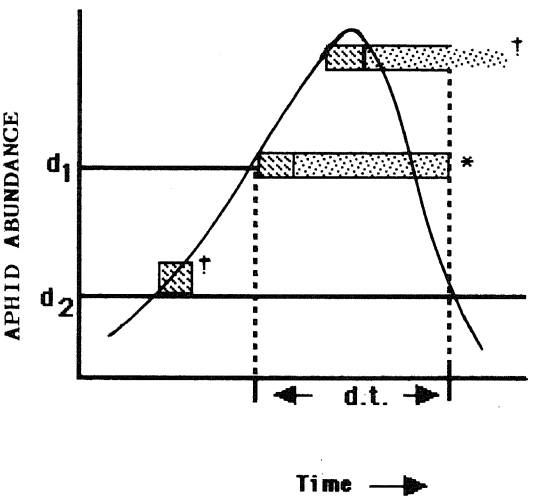


Figure 4. Reproductive behaviour of ladybirds: the period during which egg laying is compatible with the requirements for the survival of larvae. ▨ = egg laying, ⊞ = larval development (open means that larvae cannot complete their growth; †unsuccessful; *successful reproductive effort; d_1 and d_2 = threshold densities for the survival of first- and fourth-instar larvae, respectively; d.t. = developmental time) (after ref. 16).

that the optimum number of eggs laid by an adult ladybird corresponds to the maximum offspring biomass which can be supported by aphid biomass. Thus the reaction of ovipositing females to the presence of conspecific larvae in an aphid colony seems to be adaptive because it allows females (i) to choose young aphid colonies – the presence of larvae indicates that the aphid colony has been present for some time, and (ii) to avoid situations where competition for food could threaten the development of their offspring. Under these circumstances adult ladybirds are likely to leave an aphid aggregate and begin search for a new colony soon after a certain number of eggs are laid. Further studies on these aspects are currently being pursued by many investigators.

Cannibalism

An aphid colony is likely to be frequented by a succession of female ladybirds. As a result, larvae of different ages could forage on the same colony. Ladybirds which arrive late at an aphid colony are faced with optimization problem, i.e. the maximization of offspring production. Increase in predator density will bring about decrease in the number of prey available to each predator of

habitat (Hollying type II response)[17]. As a result, there will be a lower rate of increase in predators and a higher rate of cannibalism. Laboratory experiments reveal that egg and larval cannibalism are inversely related to aphid abundance (Figure 5). Conspecific eggs are more likely to be encountered and eaten than larvae because the former is immobile and laid in a cluster[18]. Field observations also suggest a density-dependent egg cannibalism in ladybirds[19]. However, ladybirds have developed effective mechanisms to avoid the risk of cannibalism: (i) the female ladybird does not oviposit in or immediately around an aphid colony because the adults and larvae aggregate there[20,21]; (ii) female ladybirds synchronize oviposition with the early development of aphid colonies, thus refraining from laying of eggs in the presence of larvae or even other females; and (iii) adult females are reluctant to eat their own eggs (Figure 6), and second-instar larvae have been seen to avoid eating sibling eggs[22]. This may be adaptive as it would reduce the chance of females eating their own eggs, whereas males are less likely to come across eggs they have sired. Non-sibling cannibalism, however, would still occur[21], and in a situation of prey scarcity this could be an effective means by which adults could fuel their reproductive efforts and eliminate potential competitors. This mechanism can ensure future availability of food for themselves and their offspring.

Predation

Ladybirds like *Adalia bipunctata* (L.), *A. decempunctata* (L.), *Coccinella septempunctata* L. and *C. undecimpunctata* L. commonly coexist in the same habitat[1]. In their coexistence their eggs and larvae are likely to occur in the same aphid colony[10]. Laboratory experiments indicate that the different species of ladybirds interact with one another, leading to predation of some of them (Figure 7)[18].

Available results suggest that interspecific predations are asymmetrical. Starving larvae and adults of *A.*

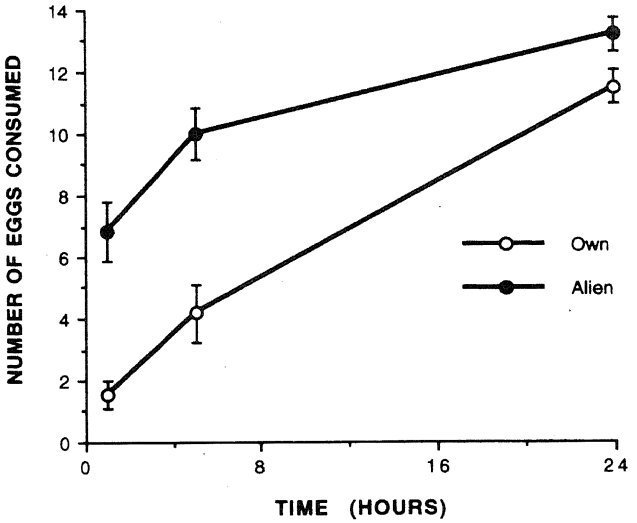


Figure 6. The mean number (± SE) of their own and alien eggs eaten by female *Adalia bipunctata* after 1, 5 and 24 h (after ref. 22).

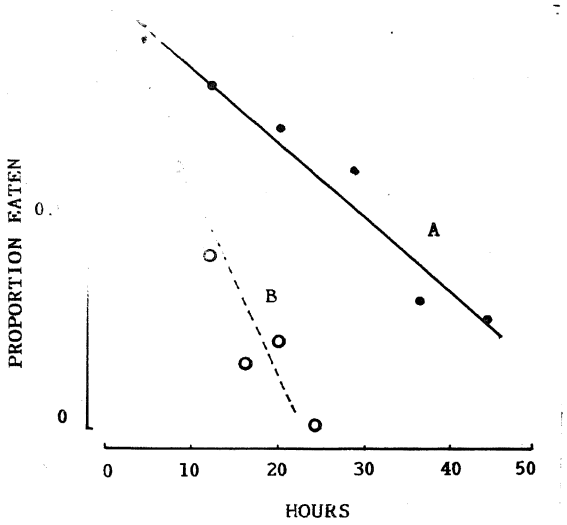


Figure 5. The proportion of the clutches of eggs (A, ●) and third-instar larvae (B, O) of *Adalia bipunctata* eaten by its fourth-instar larvae in relation to aphid density after 1, 5 and 24 h (eggs: $r^2 = 0.95$; $y = 1.007 - 0.0017x$; larvae: $r^2 = 0.91$, $y = 1.003 - 0.044x$; means of 20 observations) (after ref 18).

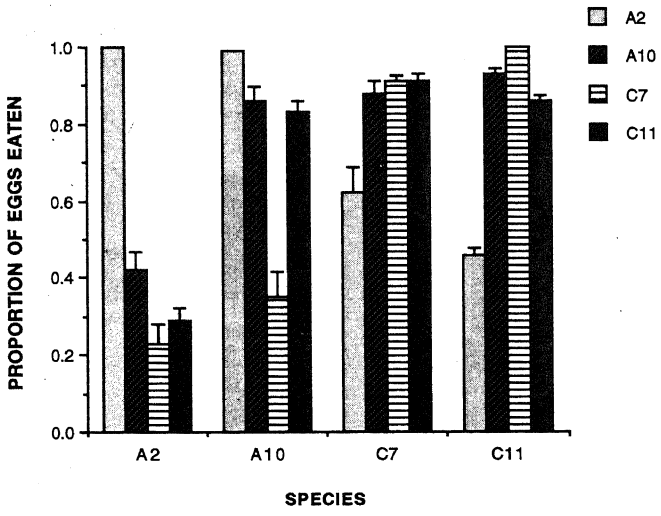


Figure 7. The proportion of eggs of their own and three other species eaten by larvae of four species of ladybirds (the four species are: *Adalia bipunctata* = A2, *A. decempunctata* = A10, *Coccinella septempunctata* = C7 and *C. undecimpunctata* = C11; 2- or 3-day-old fourth-instar larvae of the four species starved for 24 h prior to experiment; larvae were offered equal weight of eggs of the four species; *n* varies between 20 and 30) (after ref. 18).

bipunctata and *C. septempunctata* ate significantly fewer eggs of their own species when painted with a water extract of the eggs of other species. Thus, the eggs of both these species are protected to some extent from predation. The protective chemical used in this case was water-soluble (Figure 8)[18].

The results of further experiments have confirmed[19] that eating of the eggs of other species is not without risk, and that smaller species is better defended than larger species. An interesting study on the interaction between *A. bipunctata* (smaller species) and *C. septempunctata* (larger species) provides a good example. Larvae of the latter species fed on *A. bipunctata* eggs showed a significant increase in mortality than the larvae of former species fed on *C. septempunctata* eggs. The available information on interspecific predation is limited to a few of a large number of known ladybird species[1,23]. More results in future studies on a number of other species would strengthen the current thinking on interspecific predation and its possible impact on the reproductive behaviour of ladybird beetles.

That ladybirds use bitter-tasting alkaloids as a defence against predation by ants and birds[24–28] shows the chemical defence ability of the ladybird eggs. It would appear from the results that defence is based on odour rather than on colour. Recent studies suggest that clustered eggs of either *Coccinella transversalis* (Fabricius) or *Menochilus sexmaculatus* (Fabricius) are significantly less likely to be eaten by the larvae and adults of the rival species than the eggs offered singly (Agarwala, Bhattacharya and Bardhan Roy, in preparation). However, clustering of eggs does not prove to be a deterrant in cannibalism. In this context it is interesting to note that

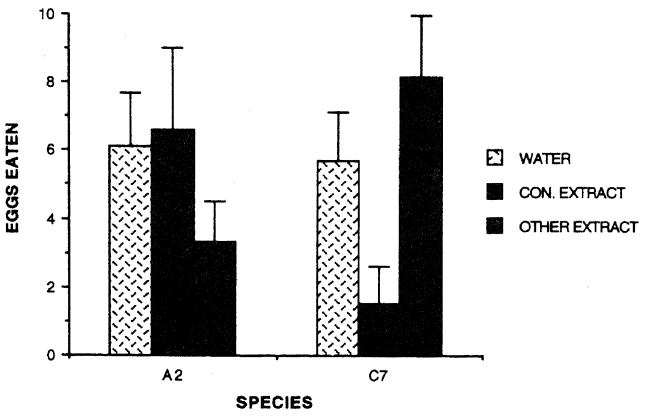


Figure 8. The effect on egg cannibalism in *Adalia bipunctata* and *Coccinella septempunctata* of painting eggs with water extract of conspecific eggs (concentrated extract) and a water extract of other species eggs (other extract) (1 mg of egg crushed in 0.15 ml of distilled water, centrifuged for 2 min and 0.05 ml of the supernatant painted on a batch of 10 eggs with a micropipette; eggs left dry for 2 h; 10 eggs offered to a starving fourth-instar larvae; experiments last 24 h; $n = 20$) (after refs 18 and 22).

alkaloids are present in the eggs, larvae, pupae and adults of ladybirds. The effectiveness of alkaloids in deterring ants is concentration-dependent[25]. Thus, clustering of eggs in ladybirds appears to be an anti-predatory device and, therefore, is another instance of adaptation for self-preservation. It would be useful to examine and define the chemical defence in ladybirds in future studies.

Conclusions

Aphid populations undergo seasonal changes in abundance. This ensures a strong selection pressure in ladybird beetles, who can successfully synchronize their development and reproduction with the phenology of aphid populations. As a result, ladybird species are able to maintain an optimum number of eggs and larvae in each aphid colony.

As aphid colony is a temporary habitat, a female ladybird chooses a young aphid colony for oviposition so that the resulting larvae could complete development. This behaviour of female ladybirds takes care of the new-born larvae, whose survival is dependent on an abundance of young aphids[29]. However, the number of eggs laid by a female has a finite relationship with prey abundance (reproductive numerical response) because (i) a satiated female is less active and, therefore, does not oviposit, and (ii) early hatching of larvae in aphid colony serves as a deterrant signal to the egg laying of females. Consequently, an optimal oviposition by a female ladybird is obvious in the early phase of aphid population development[6,14]. Such a reproductive strategy ensures that ladybird biomass production is proportional to the aphid biomass likely to be available in an aphid colony.

A lower reproductive response, thus, serves to maximize the fitness of ladybirds in terms of successful development of eggs that are laid in an aphid colony. However, a ladybird population has inherent dangers in its ecology: their positive prey-taxis in foraging and the likelihood of sharing of an aphid colony by populations of other ladybird species with a similar phenology pose considerable challenge to eggs and larvae from cannibalism and predation. Spatial distribution of ladybirds ensures that adults keep arriving at a patch of aphid colonies, resulting in predator pressure, both intraspecific and interspecific, on eggs and larvae. The pressure of intraspecific predation or cannibalism is avoided (i) in part by lower numerical response, which ensures that the optimal number of ladybird eggs laid in a particular colony should be lower, thereby not offering an opportunity of density-dependent cannibalism[19] and (ii) by the reluctance of a female to eat its own eggs, and also its inhibition in laying eggs in the presence of conspecific larvae. The ability of a ladybird to assess the quality of aphid colony diminishes the chances of egg and larval cannibalism.

Interspecific predation is subdued partly because ladybird eggs afford protection through chemical defence. Clustering of eggs in predatory ladybirds seems to be strongly adaptive. These devices not only result in increased fecundity in less time[30], so vital for its reproductive strategy, but also, more importantly, improve the chances of egg survival in two ways: (i) the defence acts as a distasteful deterrant to its predators[31,32], and (ii) some of the eggs in a cluster are inaccessible to parasites and predators[33].

The various accounts mentioned here suggest that ladybird beetles have evolved a combination of mechanisms that ensure their occurrence in an environment which suffers from limitations of temporary existence of prey and risks of predations. Some grounds of the reproductive strategy of ladybirds in response to cannibalism and predation have already been worked out but the hypothesis that clustered eggs are chemically defended against ladybird predators needs verification. Also there is a need to understand more precisely the prey–predator interactions in terms of factors which maximize the fitness of ladybird beetles in an uncertain habitat.

1. Hodek, I., *Biology of Coccinellidae*, Academic Press, Prague, 1973.
2. Ghosh, A. K., *Indian Agric.*, 1974, **18**, 1–214.
3. Blackman, R. L. and Eastop, V. F., *Aphids on the Worlds' Crop: An Identification Guide*, Wiley, Chichester, 1984.
4. Dixon, A. F. G., *Aphid Ecology*, Blackie, Glasgow, 1985.
5. Kindlmann, P. and Dixon, A. F. G., *J. Evol. Biol.*, 1992, **5**, 677–690.
6. Kindlmann, P. and Dixon, A. F. G., *Eur. J. Entomol.*, 1993, **90**, 443–450.
7. Carter, M. C. and Dixon, A. F. G., *J. Anim. Ecol.*, 1982, **51**, 865–878.
8. Kareive, P. and Odell, G., *Am. Nat.*, 1987, **130**, 232–270.
9. Agarwala, B. K. and Ghosh, A. K., *Trop. Pest Manage.*, 1988, **34**, 1–14.
10. Evans, E. W., *Oecologia*, 1991, **87**, 401–408.
11. Obrycki, J. J. and Tauber, M. J., *Ann. Entomol. Soc. Am.*, 1981, **74**, 31–36.
12. Hemptinne, J.-L. and Naisse, J., in *Ecology and Effectiveness of Aphidophaga SPB*, Academic Publishing, The Hague, 1988.
13. Hemptinne, J.-L. and Dixon, A. F. G., in *Behaviour and Impact of Aphidophaga SPB*, Academic Publishing, The Hague, 1992.
14. Hemptinne, J.-L., Dixon, A. F. G. and Coffin, J., *Oecologia*, 1992, **90**, 238–245.
15. Beddington, J. R., Hassell, M. P. and Lawton, J. H., *J. Anim. Ecol.*, 1976, **45**, 165–185.
16. Agarwala, B. K., in *Reproductive Strategy in Aphidophagous Ladybeetles (Coccinellidae) of India*, ICAR Final Technical Report, 1993
17. Hassell, Michael P., *The Dynamics of Arthropod Predator–Prey Systems*, Princeton University Press, Princeton, New Jersey, 1978.
18. Agarwala, B. K. and Dixon, A. F. G., *Ecol. Entomol.*, 1992, **17**, 303–309.
19. Mills, N. J., *Ann. Appl. Biol.*, 1982, **101**, 144–148.
20. Bank, C. J., *Anim. Behav.*, 1957, **5**, 12–24.
21. Osawa, N., *Res. Popul. Ecol.*, 1989, **31**, 153–160.
22. Agarwala, B. K. and Dixon, A. F. G., *Eur. J. Entomol.*, 1993, **90**, 45–50.
23. Majerus, M. and Kearns, P., *Ladybirds*, Richmond Publishing, England, 1989.
24. Rothschild, M., *Trans. R. Entomol. Soc. London*, 1961, **113**, 101–122.
25. Pasteel, J. M., Deroe, C., Tursch, B., Braekman, J. C., Daloze, D. and Hootele, C., *J. Insect. Physiol.*, 1973, **19**, 1771–1784.
26. Mueller, R. H., Thompson, M. E. and Dipardo, R. M., *J. Org. Chem.*, 1984, **49**, 2217–2231.
27. Rothschild, M., Moore, B. P. and Brown, M. V., *Biol. J. Linn. Soc.*, 1984, **23**, 375–380.
28. Marples, N. M., Brakfield, P. M. and Cowie, R. J., *Ecol. Entomol.*, 1989, **14**, 79–84.
29. Dixon, A. F. G., *J. Anim. Ecol.*, 1959, **28**, 259–281.
30. Courtney, S. P., *Am. Nat.*, 1984, **123**, 276–281.
31. Fisher, R. A., *The Genetical Theory of Natural Selection*, Dover, New York, 1958.
32. Stamp, N. E., *Am. Nat.*, 1980, **115**, 367–380.
33. Damman, H. and Cappuccino, N., *Ecol. Entomol.*, 1991, **16**, 163–167.

ACKNOWLEDGEMENTS. We thank the Department of Science & Technology, Govt. of India, for financing the work. We also thank Mr P. Dhanvel for help with the manuscript. We also acknowledge the criticism of the manuscript by an anonymous reviewer for improvements in the paper.

Received 8 June 1994; revised accepted 7 August 1995

Radioactivity aspects of Indian coals

M. P. Rajan, M. A. R. Iyengar and T. V. Ramachandran*

Environmental Survey Laboratory, Health Physics Division, Bhabha Atomic Research Centre, Kalpakkam 603 102, India
*Environmental Assessment Division, Bhabha Atomic Research Centre, Bombay 400 085, India

This paper presents the results of a study of the radioactivity aspects of Indian coals. The measurements were carried out by gamma spectrometry and chemical methods for estimation of U-nat, Ra-226, Ra-228, K-40, Po-210 and Pb-210 in the samples. A brief description of sampling and measurement procedures are also given. Radiation dose estimates of the population from two typical thermal power stations at Neyveli and Ennore in Tamil Nadu have been attempted using the model developed by UNSCEAR and ORNL reports as guidelines, with suitable correction for local population density, plant load factor, fly-ash release, coal consumption per day, etc.

THE global energy needs have enormously increased with rapid strides in technology and to a large extent these have been met from fossil fuels. India has to depend largely on coal reserves for her energy needs to meet the present demand and possibly in future also. The annual requirement of coal is expected to go beyond 425 million tonnes by 2000 AD compared to current annual production of about 250 million tonnes. The total energy-generating capacity in the country is expected to be about 106 GW(e) by 2000 AD. Thermal power generation, as on 1994, is about 66% of the total generating capacity, i.e. about 65000 MW(e).

Combustion of coal in thermal power stations (TPSs) gives rise to gaseous and particulate pollution in the environment, such as SO_2, NO_x, particulate matter, etc., which have long-term effects on vegetation, fish, animal life, human beings and even on buildings, structures, etc. This has also resulted in the release of large quantities of CO_2 adding to its continuous build-up of its level in the environment. It has been estimated that a 1000 MW(e) coal-fired power station consumes about 3 million tonnes of coal per year, producing 7 million tonnes of CO_2, 0.12 million tonnes of SO_2, 0.02 million tonnes of NO_x and 0.75 million tonnes of ash[1]. The present coal consumption in TPS in the country results in adding an estimated 12.2 million tonnes of fly-ash into the environment every year, of which nearly a third goes into the air and the rest is dumped on land or discharged in waterbodies.

In India, most of the coal available is of bituminous type. The quality of Indian coal is poor with 5–50% ash, 1–30% moisture, 79–82% carbon, 3.5–6.0% hydrogen, 0.8–2.4% nitrogen, 1.3–10% oxygen, 13–50% volatile matter, 0.2–2.0% sulphur and 0.01–0.4% phosphorus. Besides, coal also contains radioactivity due to primordial uranium and thorium series nuclides and K-40; during its combustion these natural radionuclides are readily released which disperse over a large area in the vicinity of TPSs[2,3]. Unlike most of the nuclear and hydroelectric power stations, coal-fired power stations in India are generally located in areas which are thickly populated and hence the environmental impact experienced by the neighbourhood population may be of significance. Due to the high ash content of coal and relatively poor fly-ash control measures, particulate pollution as well as natural radioactivity deposition due to TPSs in India are higher than in some of the western countries. However, recently built TPSs in India have installed modern pollution control technologies which mitigate the pollution levels in the environment to a considerable degree.

In the light of the possible radiological significance in the environment, studies have been carried out by several investigators in different parts of the world to assess the nature and quantity of radioactivity releases from thermal power stations using coal as fuel. Some of these studies have shown that the radiation hazard in the environment of the coal-fired power stations is comparable and sometimes even higher than that of nuclear power stations[4–6]. There is thus considerable interest in investigating the radioactivity aspects of coal and fly-ash in Indian TPSs and their significance in terms of environmental impact and public radiation exposures. This paper presents the results of a study carried out to assess the radiological aspects of certain selected TPS units in the country using both coal and lignite.

Materials and methods

Coal samples for the study were obtained from West Bengal, Chanda (Maharashtra), Singareni (Andhra Pradesh) and Talcher (Orissa) coal fields. Lignite samples were collected from various locations of the Neyveli Lignite Corporation in South Arcot District of Tamil Nadu. The fly-ash and slag (bottom ash) samples were obtained from Singareni, Ennore and Neyveli power stations. Coal, fly-ash and slag samples thus obtained were thoroughly mixed, dried and ground to a fine powder and nearly 200 g each of the sample was filled and topped up in an air-tight plastic container of 6.5 × 7.5 cm size and stored for nearly a month before

counting to ensure that nat. U and Th attain equilibrium with their respective daughters[2].

Standard sources of uranium and thorium have also been prepared in the same geometry by mixing known quantities of uranium ore and thorium nitrate salt with an inactive matrix having density close to those of the samples. For preparing K-40 standard, potassium chloride of Analar grade was used directly. Source strengths were selected in a way that would yield a low statistical error in channel counts for nearly one-hour counting. Sodium oxalate was used as a dummy matrix because it has low radioactivity content, is nonhygroscopic and has density close to the sample matrix. The overall errors due to sampling and source preparation were estimated to be ± 15% (ref. 2). The thorium nitrate and the uranium ore used for the source preparation were more than 15 years old and hence it was ensured that radium and thorium-228 had attained radioactive equilibrium with their daughters.

The selection of the 12.5 × 10.0 cm NaI (Tl) detector of Harshaw make for the present work was made keeping in view the low specific activity of the samples to be counted and the maximum energy of interest[7]. The detector offers the best compromise if the energy of interest is up to 3 MeV. Similarly, the selection of the multichannel analyser for the analysis is also based on the criteria outlined by Schonfeld[8]. The choice of the internal dimensions of the lead shield and of lining it with cadmium and copper were also made to reduce the low-energy scattered background.

Some of these samples were also analysed by radiochemical methods to determine concentrations of radionuclides like U-nat, Ra-226, Ra-228, Po-210 and Pb-210. Details regarding the radioanalytical procedures used in the study are described elsewhere[9,10].

Results

From the gamma spectrometric analysis of the samples, it was observed that the coal and its fly-ash samples contained radionuclides from both uranium and thorium series, while it was interesting to note that the lignite and fly-ash samples from Neyveli showed predominantly radionuclides from uranium series.

Tables 1 and 2 give the results of the radiochemical analysis of the samples investigated. From Table 1 it can be seen that the coal samples from different regions showed more or less similar activity content, whereas the lignite samples showed a considerable variation in the activity levels. From Table 2, it can be seen that the lignite fly-ash contains higher amount of U-nat, Ra-226, Po-210 and Pb-210 compared to the respective concentrations of these nuclides in coal fly-ash.

Among the samples investigated, the samples from Neyveli region showed evidence of higher concentration of uranium. The carbonaceous clay, which is present as a layer above the lignite seam, contains small amounts of U_3O_8 and is the main constituent which may be responsible for the observed higher concentration of uranium in the lignite samples[2,3]. Table 3 gives the

Table 1. Radioactivity content in coal and lignite samples

Source	U-nat	Ra-226	Ra-228	Po-210	Pb-210
	(Bq/kg)				
Coal					
Singareni	5.7–21.5	6.7–15.5	8.9–17.0	2.6–21.8	4.9–24.5
Bengal	22.3	14.2	17.4	17.5	16.3
Chanda	20.6	15.7	18.5	18.1	14.4
Talcher	40.3	38.9	40.2	35.8	36.1
Singrauli	53.8	54.4	45.5	50.1	48.6
Ennore	40.4–56.8	8.1–31.8	6.6–28.9	18.9–43.7	17.4–45.0
Lignite					
Neyveli	0.7-3.2	BDL-5.4	BDL-1.3	BDL-9.8	1.1–7.6
Neyveli (power station)	74.5–99.5	37.3–109.9	BDL-7.5	67.1–87.7	30.7–85.8

Table 2. Radioactivity content in fly-ash samples

Source	U-nat	Ra-226	Ra-228	Po-210	Pb-210
	(Bq/kg)				
Singareni	18.8–27.9	15.5–66.6	13.7–44.4	24.1–27.4	25.7–30.2
Ennore	73.1–121.4	16.0–94.1	15.8–79.9	5.9–55.9	7.0–69.6
Neyveli	207.5–635.3	177.5–596.4	24.4–59.4	98.9–276.4	155.8–271.4

observed levels of radionuclides in lignite from different layers as seen at the mine. The increase in the activity content of lignite fly-ash can be clearly traced to the mixing of the relatively higher activity of carbonaceous clay with lignite during the mining process and pulverization, prior to combustion in the power station.

Dose assessment

In view of the range of activities and respective levels and the wide dispersion potential of radioactivity from coal-fired TPSs, it is of interest to calculate the doses received, if any, by the public inhabiting the neighbouring environment. In our earlier publications[11,12], the radiation dose commitments to the population living within 80 km radius of thermal power stations were calculated following the 1980 UNSCEAR report. This was calculated on the basis of an assumed plant load factor of 50%, average coal consumption of 15 tons per day per MW(e) and 30% ash content of coal. A good deal of work has been done in the past few years by several investigators on assessing the doses and dose commitments from energy production in coal-fired and nuclear power plants. A review of the methods used for such calculations by De Santis[5] has been taken into account and the method described by McBride *et al.*[4] has been adopted. The method described in this report gives a clear procedure for dose assessment from a model 1000 MW(e) coal-fired and nuclear power plants.

In the present calculations the actual average plant load factor during the period of sampling, the coal consumption, and the ash content of the coal used in a particular power plant, have been used[13,14] (Table 4). These calculations were also modified in the light of ICRP recommendations of 1977, i.e. change of quality factor for alpha particles from 10 to 20 and the concept of collective effective dose equivalent commitments. The Ennore TPS near Madras is equipped with electrostatic precipitators, for which an average fly-ash control of 97% was assumed. The Neyveli TPS was equipped with mechanical or cyclone collector at the time of sample collection and a fly-ash control of 82% was assumed for this plant. These factors have been applied in deriving the dose estimates.

McBride *et al.*[4] assumed that all the daughters of the uranium and thorium series were in equilibrium and had calculated the radiation doses in the population living within 88.5 km of the plant under normal climatic conditions using the AIRDOSE computer code for the atmospheric dispersion calculations. The annual precipitation assumed in these calculations is about 900 mm. The TPSs under the present study fall in the range of 600–1000 mm annual rainfall. Hence, the data of McBride *et al.*[4] were used for calculation of dose with appropriate corrections for population density[15,16], installed capacity, coal consumption and ash content, fly-ash control factor, average load factor for the sampling period and lastly the activity release ratio of the actual activity released in Indian TPSs to that of the model TPS.

Using the above correction factors and assumption

Table 3. Radioactivity content in lignite samples of Neyveli

Sample no.	Type of sample	Bq/gm			Sample no.	Type of sample	Bq/gm		
		Ra-226	Ra-228	K-40			Ra-226	Ra-228	K-40
ASC1	Carbonaceous clay	1.20	0.07	0.37	AML1	Lignite	0.06	0.01	0.06
ASC2	Ball clay	0.02	0.02	0.02	AML2	Lignite	0.05	0.02	0.07
ASL1	Lignite	0.05	0.07	0.17	AML3	Lignite	0.05	0.03	0.08
ASL2	Lignite	0.03	0.03	0.12	ANC1	Carbonaceous clay	0.33	0.08	0.09
ASL3	Lignite	0.03	0.03	0.21	ANC2	Ball clay	0.02	0.04	0.08
AMS1	Sandstone	0.05	0.09	0.16	ANL1	Lignite	0.05	0.07	0.17
AMC1	Carbonaceous clay	0.26	0.09	0.08	ANL2	Lignite	0.04	0.03	0.12
AMC2	Ball clay	0.02	0.04	0.08	ANL3	Lignite	0.03	0.03	0.21
						Fly-ash	0.26	0.07	0.14

Table 4. Estimated dose commitments from Neyveli and Ennore thermal power stations

Name of power plant	Population density (km^{-2})	Inst. capacity (MW(e))	Ash content (%)	Avg. PLF*	Activity released (MBq/yr)			Collective effective dose equivalent commitment (man-Sv. yr^{-1})				
					Ra-226	Th-228	K-40	W.B.	Bone (1)	Lung (2)	Thyroid (3)	Total (1+2+3)
Ennore	330	330	35	28.0	518	637	1409	0.07	0.44	0.14	0.004	0.584
Neyveli	370	600	4	59.3	3684	524	262	0.36	2.49	0.60	0.022	3.112

*PLF – plant load factor.

for release considered by McBride *et al.*[4], it was found that, of the total dose received, uranium daughters contribute 75, 77, 61 and 76% of the total dose to the whole body, bone, lungs and thyroid, respectively. The corresponding figures for Ra-228, Th-228 of thorium series were 22, 18, 33 and 22% of the total dose. The contribution to the doses and dose commitments from the airborne releases of U and Th daughter products due to inhalation and direct gamma irradiation from surface deposition are 6.5, 17, 38 and 6.4% to the whole body, bone, lungs and the thyroid, respectively, the remainder arising through ingestion pathways. The doses and dose commitments from the airborne releases from thermal power stations were calculated only for the whole body, bone, lungs and thyroid as these are the doses usually considered in the context of nuclear power plants. They are also used elsewhere[12] for comparison between these two energy production technologies.

Dose estimates from K-40 have not been included because the levels of K-40 in man are homeostatically controlled. The lung is the critical organ for Rn-222 but the Rn-222 contribution to the total lung dose is only about 1 part in a million[4]. Hence, doses due to radon and its daughters are neglected in the present dose assessment and the dose assessment is done only for Th-230, Ra-226, Po-210 and Pb-210 from the U series and Th-232, Ra-228, Th-228 from the Th series.

The resulting equations based on the above considerations and assumptions for assessment of doses and dose commitments (in man-Sv. Y^{-1}) are as follows:

$$\text{Wholebody and thyroid} = 9.5 \times 10^{-8} P\ (mx + ny) \cdot d,$$
$$\text{Bone} = 2.7 \times 10^{-6} P\ (mx + ny) \cdot d,$$
$$\text{Lungs} = 7.7 \times 10^{-7} P\ (mx + ny) \cdot d,$$

where x and y refer to the radionuclides concentrations of Ra-226 and Th-228 in Bq/kg of fly-ash, d is the population density per km^2 in an 88.5 km radius environment and P is the total installed capacity in MW(e). The constants m and n are defined as follows:

m = (annual ash released from the stack in kg for 1000 MW(e)) × (dose conversion factor from U series) × (296 MBq of Ra-226)$^{-1}$,

n = (annual ash released from the stack in kg for 1000 MW(e)) × (dose conversion factor from Th series) × (185 MBq of Th-228)$^{-1}$

The values of dose conversion factors for the uranium and thorium series for different body organs have been reported earlier[17]. Plant load factors are incorporated in the computation of m and n.

The conversion of the collective dose equivalent commitments into collective effective dose equivalent commitments allows a better evaluation of the hazard of radionuclides exposures for risk estimates. Hence, all the doses estimated in the foregoing to different body organs have been converted to collective effective dose equivalent commitments and the calculated results of the dose assessments are given in detail in Table 4 for the TPSs studied.

Discussion

Among the samples analysed, the lignite and its fly-ash samples from Neyveli were found to contain more radioactivity due to uranium series radionuclides, compared to coal and its fly-ash samples from other parts of India. This gives rise to higher collective doses at Neyveli as is clear from Table 4.

The doses reported in the paper are due to the inhalation of radionuclides contained in the fly-ash released through the stack. However, for ingestion route, an investigation of environmental samples such as water, air, vegetation, fish, etc., from Ennore and Neyveli TPSs, revealed that the activities due to Ra-226, Ra-228, U-nat, etc., in these samples were in the same range as that of the samples collected from a control area free from any type of industrial pollution. Therefore, it can be concluded that the collective dose is coming essentially from the inhalation route and the dose from the ingestion route is negligible.

1. Johnes, M. S., *Atom*, 1988, **379**, 12–17.
2. Ramachandran, T. V., Ph D thesis, University of Bombay, Bombay, 1985.
3. Venkataraman, K., *J. Geol. Soc. India*, 1966, **7**, 129–133.
4. McBride, J. P., Moore, R. E., Witherspoon J. P., Blanco, R. E., *Science*, 1978, **202**, 1045–1050.
5. De Santis, V. and Longe, I., *Health Phys.*, 1984, **46**(1), 73–84.
6. Ramachandran, T. V., Lalit, B. Y. and Mishra, U. C., *Radiat. Protect. Dosim.*, 1987, **18**, 169–173.
7. Oeschger, H. and Wahlen, *Nucleonics*, 1958, **9**, 339.
8. Schonfeld, A. H., Kibbey, E. and Devies, W., Jr, Report No. DRNL-3744, 1965.
9. Lalit, B. Y., Ramachandran, T. V. and Rajan, S., *Radiat Environ. Biophys.*, 1980, **18**, 13–17.
10. Rajan, M. P., M Sc thesis, University of Bombay, Bombay, 1983.
11. Lalit, B. Y., Ramachandran, T. V. and Mishra, U. C., *Bull. Radiat. Protect.*, 1983, **6**, 125–131.
12. Mishra, U. C., Lalit, B. Y. and Ramachandran, T. V., Proceedings of the 6th IRPA Congress, Berlin (W), May 7–12, pp. 537–540.
13. Central Board of Irrigation and Power, Publ. No. 156, Malcha Marg, Chanakyapuri, New Delhi, 1982.
14. *Mon. Therm. Hydro Perform. Rev.*, Central Electricity Authority, Gen. Data Collection, CEA, Govt. of India, Sewa Bhavan, New Delhi, 1980, **8**(4), 21–23.
15. *1984 – A Reference Annual*, Chapter 1 & 27, Publications Division, Ministry of Information & Broadcasting, Publication Division, Govt. of India, New Delhi, 1984, pp. 1–13, 474–525.
16. *Statistical Abstract, India – 1982*, New Series No. 26, Central Statistical Organization, Ministry of Planning, Govt. of India, Table 12, Estimated Mid-year Populations, 1984, pp. 33–34.

17. Lalit, B. Y., Ramachandran, T. V. and Mishra, U. C., *Sci. Tot. Environ.*, 1986, **52**, 221–232.

ACKNOWLEDGEMENTS. We thank all the TPS authorities at Ennore, Neyveli, Singareni, etc., for providing us with coal and fly-ash samples and rendering help in collecting various types of samples from the station and their environment. We also thank Shri S. Krishnamony, Head, Health Physics Division; Dr K.S.V. Nambi, Head, Environmental Assessment Division; and Shri M. Sundaram, Head, PPSS, Health Physics Division, for their active interest in the study. We are also grateful to Dr N.S. Krishna Prasad, formerly of Uranium Extraction Division, BARC, for his valuable guidance and useful discussions during the course of the work.

Received 31 July 1995; revised accepted 28 August 1995

Composition and enzyme activities in *Aspergillus flavipes* grown on crude petroleum oil and glucose

J. Savitha* and C. V. Subramanian**

Centre for Advanced Study in Botany, University of Madras, Guindy Campus, Madras 600 025, India
*Present address: Department of Microbiology, Bangalore University, Bangalore 560 056, India
**'Anjaneya', 62, Dr Ramaswamy Salai, West K. K. Nagar, Madras 600 078, India

Aspects of the growth, cellular composition and enzymology of *Aspergillus flavipes* grown on a medium with crude oil or glucose as the sole source of carbon are reported. When grown on crude oil, the total lipid content of *A. flavipes* was considerably greater than on glucose medium. A large number of fatty acids, particularly saturated fatty acids, were observed when grown on crude oil and more lipid was excreted into the growth medium. The total mycelial carbohydrate and glycogen contents were considerably less in crude oil medium, whereas the mycelial protein content was similar in both media. The activity of some enzymes like NAD^+-alcohol dehydrogenase, $NADP^+$-aldehyde dehydrogenase, iso-citrate lyase and catalase was greater when crude oil was the carbon source. Conversely, the activity of succinate dehydrogenase was greater after growth on glucose medium.

THE subject of degradation and utilization of hydrocarbons by fungi has received considerable attention[1–3] but much more remains to be done[4]. A strain of *Aspergillus flavipes* isolated from the effluents of the Madras Oil Refineries that can degrade hydrocarbons in crude petroleum oil was reported earlier[5]. Work on utilization of hydrocarbons has mainly been carried out on a few selected bacteria, and some fungi, notably species of *Candida* and the kerosene fungus *Cladosporium resinae*[6–8]. Much of this work has used specific hydrocarbons such as alkanes of known chain length. In the present study, however, we have examined the potential of *A. flavipes* isolated from the effluents of the Madras Oil Refineries to degrade crude petroleum oil, which often forms around 2% (v/v) or more of the effluent from the refinery. The emphasis in the study has been on the utilization of crude oil *per se* rather than particular components of the oil.

We have examined the amount of total protein, total carbohydrate, glycogen, total lipids, total sterol, fatty acids (saturated and unsaturated) in *A. flavipes* grown with either crude oil or glucose as the sole source of carbon. As a preliminary to understanding the physiology of utilization of crude oil, some data are given on the enzymology of crude oil and glucose utilization.

Materials and methods

Organism. A single spore isolate of *A. flavipes* isolated from the effluents of the Madras Oil Refineries Ltd., Manali, Madras, was used throughout the study. It was maintained on potato dextrose agar (PDA).

Culture techniques. Standard microbiological techniques were used in the preparation of media, sterilization etc.[9]. The fungus was grown in a mineral salts solution of Bushnell and Haas[10] supplemented with 0.1% (w/v) yeast extract, adjusted to pH 6.0. Glucose (10 g l^{-1}) was added to mineral salts solution before the autoclaving. Crude oil was filter-sterilized (pore size, 0.41 μm) and added to sterile mineral salts solution to give 10 g l^{-1}. Flasks (250 ml) containing 100 ml medium were inocu-

lated with 1 ml of a spore suspension (1×10^6 spores ml^{-1}) and incubated at 28 ± 2°C under static conditions. Spores were harvested in sterile distilled water from PDA plates overlaid with cellophane.

Determination of fungal growth. The mycelial mat was harvested by filtering through a predried and weighed Whatman no. 1 filter paper. The mat was washed three times with chloroform and then with distilled water, and dried at 100°C for 25 h; the dry weight of the mycelium was then determined.

Protein estimation. Total protein was estimated by the method of Lowry *et al.*[11] using bovine serum albumin as standard.

Determination of total sugar and glycogen. Total sugar and glycogen were estimated according to Roe[12] and Morales *et al.*[13], respectively.

Estimation of total lipid. Total lipid was extracted from 10 g fresh weight of mycelium as described by Folch *et al.*[14] and estimated gravimetrically.

Separation and estimation of neutral, glycolipid and phospholipid by column chromatography. The total lipid was fractionated into its constituent lipid classes by silicic acid column chromatography[15]. Neutral lipid was eluted with eight column volumes of chloroform, glycolipid with one column volume of 99% (v/v) acetone and phospholipid with eight column volumes of methanol. Solvents were evaporated and the lipid fractions were estimated gravimetrically. The percentage of lipid fractions in total lipid was then calculated.

Extraction of extracellular lipid. Extracellular lipid was extracted from the culture medium by the method of Siporin and Cooney[16]. The mycelial mat was removed from the medium by filtering the culture through two layers of Whatman no. 1 filter paper. In cultures grown on crude oil a thin layer of residual oil was present on the surface of the medium. This layer was removed by shaking 100 ml of medium with 10 ml petroleum ether and removing the upper solvent phase. The aqueous phase was adjusted to pH 2 with HCl and extracted three times with two volumes of chloroform : methanol (2 : 1 v/v).

Extraction of total extracellular lipid by thin layer chromatography (tlc). The total extracellular lipid was separated into free fatty acid and phospholipid fractions by tlc using chloroform : acetone : methanol : acetic acid : water (8 : 1 : 5 : 25 : 0.25 v/v) as solvent[16]. Lipids were visualized by iodine vapour. Spots corresponding to standard free fatty acid and phospholipid were eluted with hexane, dried in nitrogen atmosphere and stored at –4°C. The amount of total free fatty acid[17] and phospholipid[18] was estimated spectrophotometrically.

Fatty acid analysis. Fatty acids were converted to their methyl esters using borontrifluoride : methanol[19] and separated by gas chromatography using a polar column (10% w/w, diethylene glycol succinate on chromosorb/Aw) operated at 190°C. The relative proportions of the different fatty acids were calculated by manual measurement.

Measurement of enzyme activity. Activities of the NAD^+-alcohol dehydrogenase, $NADP^+$-alcohol dehydrogenase and $NAD(P)^+$-aldehyde dehydrogenase were measured spectrophotometrically at 340 nm following the procedures of Bonnichsen[20], Demoss[21] and Seegmiller[22], respectively. Succinate dehydrogenase was assayed[23] at 428 nm. Isocitrate lyase was measured[24] at 324 nm. Catalase was measured as described by Beers and Sizer[25].

Statistical analysis. All experiments were carried out in duplicate and each experiment repeated three times, except the cellular and extracellular fatty acid experiments, which were repeated twice with no replicates. Data obtained were analysed statistically[26]. Student's *t* test was used to calculate the degree of significance between glucose and crude oil.

Results

Growth and composition (cellular) of A. flavipes. When grown in static culture on glucose the stationary phase was reached after 6 days, but on crude oil the growth was not complete until the twelfth day. To determine the cellular composition, the fungus was harvested at four sample times during its growth period (Figure 1).

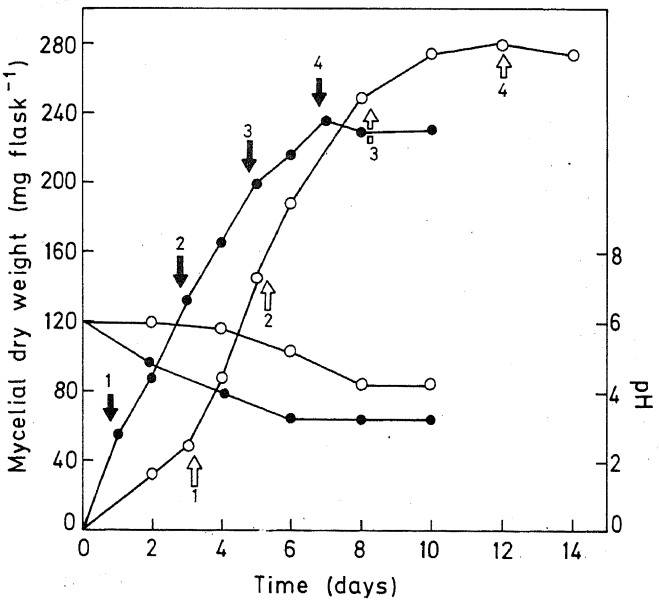


Figure 1. Growth of *A. flavipes* grown on medium with glucose and crude petroleum oil as the sole source of carbon: ● glucose; O crude oil. Arrows indicate sample times. The experiment was carried out in triplicate.

Variations in cellular composition of *A. flavipes* were observed after growth on crude oil or glucose media. Total carbohydrate and glycogen contents were approximately twofold greater in mycelia grown on glucose although protein content was not significantly different between the two media (Table 1). Considerable increases in total lipid and sterol contents were observed for the crude-oil-grown cultures (Table 1). In cultures grown on crude oil, neutral lipid was the major fraction followed by glycolipid, whereas in glucose-grown cultures glycolipid was the major fraction, followed by neutral lipid. However, on both media phospholipid was the minor fraction (Table 2). The lipid of crude-oil-grown cultures contained a variety of fatty acids compared to that of glucose-grown cultures. Also, the crude-oil-grown cultures showed an increase in the percentage of total unsaturated fatty acid with time (Table 3).

Extracellular lipid. In the presence of crude oil

Table 1. Cellular composition of *A. flavipes* grown on medium with glucose and crude oil as the sole source of carbon

Sample time	Total protein[a]		Total carbohydrate[a]		Total glycogen[a]		Total lipid[a]		Total sterol[a]	
	Glucose	Crude oil	Glucose	Crude oil	Glucose	Crude oil	Glucose	Crude oil	Glucose	Crude oil
1	180 ± 5	190 ± 9	191 ± 6***	76 ± 5	76 ± 3	32 ± 2	26 ± 2	76 ± 5***	14 ± 1	52 ± 3***
2	100 ± 6	106 ± 6	285 ± 15***	105 ± 9	163 ± 8	67 ± 5	41 ± 3	264 ± 10***	24 ± 2	104 ± 5***
3	86 ± 5	91 ± 4	309 ± 15***	143 ± 11	242 ± 18	90 ± 5	37 ± 3	219 ± 6***	28 ± 1	114 ± 6***
4	143 ± 7	147 ± 12	278 ± 13***	152 ± 9	233 ± 15	99 ± 4	22 ± 2	142 ± 2***	18 ± 1	92 ± 5***

Each value represents the mean ± SEM of six estimations.
[a]Values expressed as µg/mg dry weight.
***$P < 0.001$.

Table 2. Composition of total cellular lipids of *A. flavipes* grown on medium with glucose and crude oil as the sole source of carbon

Sample time	Total lipid					
	Neutral lipid[†]		Glycolipid[†]		Phospholipid[†]	
	Glucose	Crude oil	Glucose	Crude oil	Glucose	Crude oil
1	28 ± 3	77 ± 5**	60 ± 2**	18 ± 0.9	12 ± 0.8**	3 ± 0.01
2	30 ± 1	83 ± 3**	67 ± 0.8**	13 ± 0.4	4 ± 0.1	2 ± 0.1
3	45 ± 3	75 ± 1**	52 ± 3**	22 ± 1	5 ± 0.08	4 ± 0.1
4	41 ± 2	72 ± 2**	52 ± 3**	25 ± 0.2	2 ± 0.2	4 ± 0.2

Each value represents the mean ± SEM of six estimations.
[†]Percentage of total lipid.
**$P < 0.01$.

Table 3. Fatty acid composition of total cellular lipid of *A. flavipes* grown on medium with glucose and crude oil as the sole source of carbon

Carbon source	Sample time	Fatty acid*											
		Saturated									Unsaturated		
		10:0	12:0	13:0	14:0	15:0	16:0	17:0	18:0	Total[†]	16:1	18:1	Total[§]
Glucose	1	0	0	0	tr	0	37.5	0	0	37.5	0	62.4	62.4
	2	0	0	0	tr	0	33.9	0	0	33.9	0	66.0	66.0
	3	0	tr	0	tr	0	0	0	45.5	45.4	0	54.5	54.5
	4	0	tr	0	tr	0	0	0	33.3	33.3	0	66.6	66.6
Crude oil	1	1.6	2.3	9.6	8.4	7.5	3.4	tr	23.9	56.8	26.6	16.5	43.1
	2	1.6	9.1	0	6.7	12.7	6.1	tr	6.9	43.3	18.0	38.6	56.6
	3	0	0	0	0	5.1	4.6	tr	4.6	14.9	21.5	63.8	85.4
	4	0	0	0	0	1.4	3.3	tr	4.3	9.1	12.2	78.5	90.8

*Recorded as percentage of total fatty acid: tr – trace.
[†]Total saturated fatty acid; [§]Total unsaturated fatty acid.
10:0, Decanoic acid; 12:0, dodecanoic acid; 13:0, tridecanoic acid; 14:0, tetradecanoic acid; 15:0, pentadecanoic acid; 16:0, hexadecanoic acid; 17:0, heptadecanoic acid; 16:1, *cis*-Δ9-hexadecanoic acid; 18:1 *cis*-9-octadecanoic acid.

Table 4.

a, Total extracellular lipid† in medium with glucose and crude oil as the sole source of carbon following growth of *A. flavipes*

Sample time	Glucose	Crude oil
1	0	24 ± 0.9***
2	25 ± 0.4	71 ± 2.3***
3	28 ± 0.9	86 ± 4.5***
4	27 ± 1.1	104 ± 6.9***

†Values expressed as mg/l of medium.
***$P < 0.001$.

b, Composition of extracellular lipid† in medium with glucose and crude oil as the sole source of carbon following growth of *A. flavipes*

	Free fatty acid		Phospholipid	
Sample time	Glucose	Crude oil	Glucose	Crude oil
1	0	66 ± 2***	0	14 ± 0.2
2	40 ± 2	77 ± 4**	19 ± 0.7	12 ± 0.1
3	57 ± 1	79 ± 4**	17 ± 0.3	11 ± 0.3
4	56 ± 2	79 ± 4**	17 ± 0.4	9 ± 0.2

Each value represents the mean ± SEM of six estimations.
†Recorded as percentage of total lipid.
***$P < 0.001$,
**$P < 0.01$.

A. flavipes excreted three to four times more lipid into the growth medium than when grown on glucose (Table 4 *a*). When the total lipid was fractionated by column chromatography, it failed to separate into definite fractions such as neutral lipid or glycolipid. However, after tlc free fatty acids were found to be the major lipid fraction. The free fatty acid content was nearly double in the medium with crude oil than in the medium with glucose (Table 4 *b*). The level of phospholipid was significantly higher in glucose medium than in crude oil medium. Methyl esters prepared from free fatty acid fractions of lipid extracted from medium with glucose or crude oil gave both saturated and unsaturated even-chain fatty acids; no odd-chain fatty acids were detected (Table 5).

Enzymology. In cultures grown on crude oil, the activities of NAD^+-alcohol dehydrogenase and $NADP^+$-aldehyde dehydrogenase were nearly three and two times more (respectively) than that of glucose-grown cultures, whereas $NADP^+$-alcohol dehydrogenase and NAD^+-aldehyde dehydrogenase did not show any significant difference in their activity on the two media (Table 6). Succinate dehydrogenase activity was considerably lower in the crude-oil-grown cultures, whereas isocitrate lyase and catalase activities were considerably higher in the crude-oil-grown cultures (Table 6).

Table 5. Fatty acid composition† of extracellular lipid excreted by *A. flavipes* in medium with glucose and crude oil as the sole source of carbon

Carbon source	Sample time	Saturated fatty acid 12:0	16:0	18:0	Total saturated fatty acid	Unsaturated fatty acid 18:1
Glucose	1	tr	tr	tr	tr	tr
	2	tr	24.1	35.8	60.0	40.0
	3	tr	26.2	39.3	66.1	33.8
	4	tr	32.0	28.1	60.1	25.0
Crude oil	1	19.0	33.2	tr	52.2	47.7
	2	31.4	33.4	tr	64.8	35.1
	3	9.1	27.2	30.0	66.3	33.6
	4	25.1	14.8	8.6	48.6	51.3

tr – trace.
†Each value recorded as percentage of total lipid.
12:0, Dodecanoic acid; 16:0, hexadecanoic acid; 18:0, octadecanoic acid; 18:1, *cis*-Δ9-octadecanoic acid.

Table 6. Activity† of some enzymes of *A. flavipes* grown on medium with glucose and crude oil as the sole source of carbon

Sample time	NAD^+-alcohol dehydrogenase		$NADP^+$-alcohol dehydrogenase		NAD^+-aldehyde dehydrogenase		$NADP^+$-aldehyde dehydrogenase		Succinic dehydrogenase		*iso*-Citrate lyase		Catalase	
	Glucose	Crude oil	Glucose	Crude oil	Glucose	Crude oil	Glucose	Crude oil	Glucose	Crude oil	Glucose	Crude oil	Glucose	Crude oil
1	21.6±1.4	32.5±2.9*	15.3±0.6	18.1±0.4	22.1±0.2	20.2±0.1	40.2±1.1	41.7±2.1**	22.0±1.2	14.0±1.4	8.1±0.7	23.7±1.4***	22.1±2.2	140.4±6.7***
2	19.3±0.6	51.3±1.8***	17.2±0.7	17.2±0.2	25.5±0.1	21.7±0.5	35.8±3.2	61.2±4.3**	41.4±3.2	21.0±1.4	17.4±1.3	48.6±2.8***	70.0±2.6	293.6±4.3***
3	20.0±0.9	61.0±1.9***	16.2±0.5	18.8±0.8	27.2±0.5	29.5±0.4	30.8±1.5	62.5±5.8**	54.8±4.3	13.8±0.9	12.2±1.7	56.5±4.0***	53.2±1.8	244.7±10.0***
4	16.4±1.5	52.2±3.9***	19.7±0.2	16.1±0.2	30.2±0.8	27.2±0.8	23.8±1.7	53.7±3.0**	40.2±2.3	10.2±0.6	18.4±0.8	40.6±2.0***	21.2±0.5	120.4±6.2***

†Activity expressed as nmol/min/mg protein.
***$P < 0.001$.
**$P < 0.01$.
*$P < 0.05$.
Each value represents the mean ± SEM of six estimations.

Discussion

It is interesting to note that the lipid content of *A. flavipes* was considerably greater (sixfold) when grown on crude oil as the sole carbon source compared to glucose. *A. flavipes* accumulates lipid more efficiently than either *Candida lipolytica*[29] or *Cladosporium resinae*[30] where two to threefold increases in lipid content of cultures grown on *n*-hexadecane have been reported. Growth of *A. flavipes* on crude oil resulted in the occurrence of a greater number of fatty acids, particularly saturated fatty acids (Table 3). The fatty acids are oxidation products of *n*-alkanes in the crude oil which are then incorporated into the cellular lipids. The content of saturated fatty acids decreases progressively in cultures grown on crude oil. This may be due to some of the short-chain saturated fatty acids being converted into long-chain fatty acids, especially unsaturated fatty acids, which can be used to produce storage lipids[4].

The function of extracellular free fatty acids, which increase during growth on crude oil (Table 4 *b*), is possibly to emulsify the hydrocarbons to facilitate their uptake by the fungus[26].

Induction of NAD^+-alcohol dehydrogenase and $NADP^+$-aldehyde dehydrogenase in crude-oil-grown *A. flavipes* suggests their involvement in the oxidation of alkanes in crude oil. Iso-citrate lyase is the key enzyme of the glyoxylate cycle which plays an important role in the biosynthesis of cellular components such as malate or succinate. Activity of isocitrate lyase was significantly greater when *A. flavipes* was grown on crude oil. Conversely, the activity of succinate dehydrogenase was greater in glucose medium (Table 6). From these results, it is possible to infer that the intermediates of tricarboxylic acid cycle are synthesized from acetyl COA derived from the degradation of fatty acids via glyoxylate cycle[29]. Peroxisomes are thought to contain several enzymes involved in fatty acid metabolism such as catalase, iso-citrate lyase, etc. The increase in these enzymes observed in crude-oil-grown *A. flavipes* suggests an increase in the number of peroxisomes although we have no data to support this supposition. However, this finding is in agreement with those of other workers, who report a similar increase in catalase and iso-citrate lyase activities and increase in the number of peroxisome for *C. resinae*[4] and *Candida tropicalis*[30].

The work described here is the first report on the composition and enzyme activities in a fungus grown on crude petroleum oil and glucose. The most significant aspect of the study lies in the fact that an attempt has been made here to elucidate the metabolism and enzymology of *A. flavipes* when grown on crude petroleum oil, which is a natural component of industrial wastes and effluents.

1. Cofone, L., Walker, J. D. and Cooney, J. J., *J. Gen. Microbiol.*, 1973, **12**, 126–134.
2. Davies, J. S. and Westlake, D. W. S., *Can. J. Microbiol.*, 1979, **25**, 146–156.
3. Sheridan, J. E. and Soteros, J. J., *Int. Biodeterior. Bull.*, 1974, **10**, 105–107.
4. Ratledge, C., in *Development in Biodegradation of Hydrocarbons* (ed. Watkinson, R. J.), Applied Science Publishers Ltd., London, 1978, vol. 1, pp. 1–46.
5. Savitha, J., *Curr. Sci.*, 1986, **55**, 1248–1249.
6. Cooney, J. J., Siporin, C. and Smucker, R. A., *Bot. Mar.*, 1980, **23**, 227–232.
7. Cooney, J. J. and Proby, C. M., *J. Bacteriol.*, 1971, **108**, 777–781.
8. Osumi, M., Eukuzumi, F., Teranishi, Y., Tanaka, A. and Fukui, S., *Arch. Microbiol.*, 1975, **103**, 1–11.
9. Booth, C., *Methods Microbiol.*, 1971, **4**, 1–47.
10. Bushnell, L. D. and Haas, H. F., *J. Bacteriol.*, 1941, **41**, 653–673.
11. Lowry, O. H., Rosenbrough, N. J., Farr, A. L. and Randall, R. J., *J. Biol. Chem.*, 1957, **226**, 497–509.
12. Roe, J. H., *J. Biol. Chem.*, 1955, **212**, 335–343.
13. Morales, M. A., Jabbagy, A. J. and Terenzi, H. P., *Neurospora Newslett.*, 1973, **20**, 24–45.
14. Folch, J., Less, M. and Sloane-Stanley, G. H., *J. Biol. Chem.*, 1957, **226**, 497–509.
15. Rouser, G., Kritcheuski, G., Simon, G. and Nelson, G. J., *Lipids*, 1967, **2**, 37–40.
16. Siporin, C. and Cooney, J. J., *Appl. Microbiol.*, 1975, **29**, 604–609.
17. Chromy, V., Gergel, J., Vonznicek, J., Krombholzova, L. and Muril, J., *Clin. Chim. Acta*, 1977, **80**, 327–332.
18. Barlett, G. R., *J. Biol. Chem.*, 1959, **234**, 446–468.
19. Morrison, W. R. and Smith, L. M., *J. Lipid. Res.*, 1964, **5**, 600–608.
20. Bonnichsen, R. K., *Acta Chem. Scand.*, 1950, **4**, 715.
21. Demoss, R. D., *J. Bacteriol.*, 1954, **68**, 252–257.
22. Seegmiller, J. E., *J. Biol. Chem.*, 1952, **201**, 629–630.
23. Steine, G. J., *Can. J. Microbiol.*, 1967, **13**, 1203–1205.
24. Dixon, G. H. and Kornberg, H. L., *Biochem. J.*, 1959, **72**, 3.
25. Beers, R. F. and Sizer, I. W., *J. Biol. Chem.*, 1952, **195**, 133–340.
26. Panse, V. G. and Sukhatme, P. V., *Statistical Methods for Agricultural Workers*, Indian Council of Agricultural Research Publishers, New Delhi, 1967, p. 381.
27. Nyns, E. J., Chiang, N. and Wiaux, A. L., *Antonie van Leeuwenhoek*, 1968, **34**, 197–204.
28. Walker, J. D. and Cooney, J. J., *J. Bacteriol.*, 1973, **115**, 635–639.
29. Makula, R. A and Finnerty, W. P., *J. Bacteriol.*, 1972, **112**, 398–407.
30. Tanaka, A. and Ueda, M., *Mycol. Res.*, 1993, **97**, 1025–1044.

ACKNOWLEDGEMENTS. One of the authors (JS) acknowledges UGC for the award of Junior Research Fellowship in the project 'Mycoflora of Industrial Wastes and Effluents'. We thank Dr S. K. Nigam, Head, Pharmaceutical Chemistry Lab., National Botanical Research Institute, Lucknow, for carrying out the fatty acid analysis.

Received 22 March 1995; revised accepted 2 September 1995

Liv.52 pretreatment inhibits the radiation-induced lipid peroxidation in mouse liver

N. G. Ganapathi and Ganesh Chandra Jagetia

Department of Radiobiology, Kasturba Medical College, Manipal 576 119, India

The effect of Liv.52 on mice liver mitochondrial lipid peroxidation was measured in terms of malanoldehyde (MDA) generation after exposure to 3.0 Gy of gamma radiation at ¼, ½, 1, 2, 4, 8, 24 and 48 h post-irradiation. The lipid peroxidation level increased significantly from ¼ h and reached a peak level at 4 h post-irradiation, and gradually declined thereafter in the irradiated control group. Liv.52 treatment before irradiation inhibited the formation of lipid peroxidation significantly at all the post-irradiation sampling time.

LIPID peroxidation is initiated by reactive oxygen species (ROS) in aerobic living organisms. Superoxide radical (O_2^-) and hydrogen peroxide (H_2O_2) are generated during metabolism, oxidation and irradiation, and they in turn are converted to hydroxyl free radical ($OH^\cdot$). The hydrogen atom in unsaturated fatty acids can be abstracted by the OH radical following the formation of lipid alkyl radical ($R^\cdot$), which initiates the chain reaction of lipid peroxidation in an aerobic condition.

Lipid peroxidation alters the fluidity of biological membranes and then causes cell degradation, affecting the biological defence mechanism. The product of lipid peroxidation, such as malanoldehyde (MDA), damages the enzyme system and DNA[1]. Lipid peroxidation has been used as an end point to study the action of oxidizing and free-radical-producing agents as well as to investigate the effects of intracellular radical scavengers (i.e. as an overall measurement of oxidative injury to the cellular membranes). Lipid peroxidation is considered to be an important effect of ionizing radiation on biological membranes[2–4].

Radiation-induced lipid peroxidation has been reported to be caused by superoxide radicals[5]. However, later studies confirmed that the hydroxyl radical is the most active species involved in radiation-induced lipid peroxidation[3,6–8].

Natural products such as herbal medicines have only recently begun to receive some attention as possible modifiers of the radiation response. Liv.52, a non-toxic herbal preparation composed of *Capparis spinosa*, *Cichorium intybus*, *Solanum nigrum*, *Cassia occidentalis*, *Terminalia arjuna*, *Achillea millefolium* and *Tamarix gallica*, has been reported to be clinically effective in treating hepatotoxicity and a wide range of hepatic disorders[9–11]. The radioprotective potential of Liv.52 was demonstrated for the first time in mice[12]. Recently, it has been reported that prior administration of Liv.52 protected mouse bone marrow against radiation-induced chromosome damage[13,14]. A significant decrease in lipid peroxidation, followed by an increase in the tocopherol level in Liv.52-fed groups, was observed[15]. Saxena and Garg[16] also reported that the feeding of Liv.52 to rats gave a marked protection against the increase in lipid peroxidation and phospholipid contents of microsomal fraction.

Therefore, it was desired to obtain an insight into the effect of Liv.52 on lipid peroxidation in mouse liver exposed to 3 Gy of ^{60}Co gamma radiation.

Six- to eight-week-old male Swiss albino mice weighing 25–30 g were selected from an inbred colony maintained under controlled conditions of temperature (23 ± 2°C), humidity (50 ± 5%) and light (10 and 14 h of light and dark, respectively). The animals were given sterile food prepared as per the standard formulation (wheat 70%, Bengal gram 20%, fish meal 5%, yeast powder 4%, sesame oil 0.75% and shark liver oil 0.25%) in the laboratory and water *ad libitum*. Throughout the experiment 5–6 animals were housed in a sterile polypropylene cage containing sterile paddy husk (procured locally) as bedding material. The irradiation was carried out using telecobalt therapy source (Gammatron, Siemens, Germany).

A batch of 10 immobilized animals (achieved by inserting cotton plugs in the restrainer) was kept prone in specially designed well-ventilated acrylic box during exposure. The irradiation was carried out at a dose rate of 0.88 Gy/min. The source-to-animal distance was 60–63 cm. The exposure area of 20 × 20 cm was also kept constant. The dosimetry calculation were done by Dr J. G. R. Solomon, Department of Radiotherapy and Oncology, K. M. C. Manipal.

As the Liv.52 powder is insoluble in water, a dose of 500 mg/kg body wt of Liv.52 powder was suspended in 10 ml of 5% dextrose solution containing gum acacia. 0.01 ml/g body wt of Liv.52 or 5% dextrose was administered orally once a day for 7 days using 22-gauge oral feeding needles (Popper & Sons Inc., New York, USA).

The animals were divided into two groups. One group of animals was fed with 5% dextrose solution and served as the control group, while the other group received 500 mg/kg body wt of Liv.52 powder in 5% dextrose solution. After 1 h of administration on day 7, the animals of both groups were exposed to 3 Gy of ^{60}Co gamma radiation. Subsequently, a few animals were also treated with 5% dextrose and Liv.52 but without irradiation, for the sake of comparison. Usually, five animals

Liv.52 is a commercial product marketed by The Himalaya Drug Co., Bangalore. Use of the trade name does not imply endorsement of the product.

were used for each time period for each group studied. A total of 90 animals were used for the whole experiment.

The animals from both groups were anaesthetized by administration of overdose of diethyl ether anaesthesia and were killed at ¼, ½, 1, 2, 4, 8, 24 and 48 h post-irradiation. The livers of the animals were perfused with isotonic saline and removed and weighed. A 10% liver homogenate was prepared in 0.2 M Tris–HCl buffer (pH 7.4) using Yamato LSG homogenizer LH-21 (Japan). The lipid peroxidation in liver was estimated by the method of Konings and Drijver[17]. Briefly, the homogenate was centrifuged at 50,000 × *g* (Sorvall, RC 5B, USA) for 10 min at 4°C and the pellet was discarded. The resultant supernatant was centrifuged again for 60 min at 4°C to obtain mitochondrial pellet. A homogeneous suspension of mitochondria was prepared in 0.2 M Tris–HCl buffer (pH 7.4). To this suspension 150 mM KCl, 0.3 mM ascorbic acid and 0.2 M Tris–HCl buffer (pH 7.4) were added, and it was incubated at 37°C for 1 h. After incubation, trichloroacetic acid was added followed by the addition of 0.67% TBA. The test tubes were kept in boiling water bath for 15 min. The tubes were removed from the water bath and allowed to cool. The tubes were centrifuged and the resultant supernatant was filtered through Whatman filter paper. The absorbance of the sample was read against the blank at A_{max} 532 nm in a UV–VIS recording spectrophotometer (Shimadzu UV–VIS 260, Japan). Mitochondrial protein concentration was determined by the method of Bradford[18]. The data were analysed statistically by Student's *t* test on an IBM/PC.

The lipid peroxidation was expressed in terms of nmol MDA/mg protein. Liv.52 treatment did not alter the level of lipid peroxidation (Table 1).

The exposure of animals to 3.0 Gy ^{60}Co gamma radiation resulted in a significant increase in lipid peroxidation at all the time periods studied compared to the sham-irradiated control. The activity of lipid peroxidation started increasing at ¼ h and reached a peak level by 4 h post-irradiation (Figure 1). This increase was approximately 1.88, 2.4, 2.0, 2.6 and 3 folds higher for ¼, ½, 1, 2, and 4 h post-exposure, respectively, compared to the sham-irradiated animals. An abrupt decline in the lipid peroxidation level was observed at 8 h, which was approximately half of the preceding sampling time. However, lipid peroxidation did not attain normal level even by 48 h post-irradiation (Table 1).

In the Liv.52 + irradiated group a significant increase in lipid peroxidation compared to sham-irradiated animals was observed only at ½ h post-irradiation. However, Liv.52 administration before irradiation inhibited lipid peroxidation significantly at all the time periods studied. It was 23.61, 21.64, 24.93, 43.99, 52.04, 26.81, 9.21 and 20.81% lower for ¼, ½, 2, 4, 8, 24 and 48 h post-irradiation, respectively, compared to the concurrent irradiated controls.

Lipid peroxidation induced by radiation is known to be due to the attack of free radicals on the fatty acid component of membrane lipids[7,19,20]. Mitochondrial membranes contain high percentage of polyunsaturated fatty acids and are, therefore, susceptible to free-radical attack[21]. Damage of mitochondrial structures and enhanced lipid peroxidation by ionizing radiation has been reported[22,23]. Lipid peroxidation also results in mitochondrial swelling and disintegration[24].

Irradiation of animals to 3.0 Gy of gamma radiation resulted in a significant increase in lipid peroxidation level at all the time intervals compared to the sham-irradiated group. The lipid peroxidation level started increasing at ¼ h and reached a peak level at 4 h after exposure. This increase was approximately 1.88, 2.4, 2, 2.6 and 3 folds higher for ¼, ½, 1, 2, and 4 h post-irradiation respectively, compared to the sham-irradiated group. The glutathione depletion was maximum at 15 and 45 min after exposure to 3.0 Gy of gamma radiation (unpublished data). These findings are in good agreement with the earlier findings, where maximum depletion of GSH at 15–30 min and the highest lipid peroxidation 4 h after 3.0 Gy of gamma radiation have been reported[25]. Similarly, a maximum depletion in GSH at 15–30 min and lipid peroxidation as well as liver necrosis at 2–4 h after administration of allyl alcohol in mice have been observed[26,27]. Lipid peroxidation has been suggested as one of the main causes of radiation-induced membrane damage[2,3]. The lipid peroxidation has been found to increase with increase in radiation dose in rat liver microsomes and in mice erythrocytes[28,29].

Table 1. Effect of Liv.52 on the radiation-induced lipid peroxidation in mouse liver

Time (min/h)	nmol MDA/mg protein (mean ± SEM)	
	3.0 Gy	Liv.52 + 3.0 Gy
¼	3.43 ± 0.22^a	2.62 ± 0.09^d
½	4.62 ± 0.20^a	3.62 ± 0.14^d
1	3.89 ± 0.13^a	2.92 ± 0.23^e
2	5.16 ± 0.10^a	2.89 ± 0.04^a
4	6.11 ± 0.23^a	2.93 ± 0.31^e
8	3.99 ± 0.10^a	2.92 ± 0.09^c
24	3.15 ± 0.06^b	2.86 ± 0.06^f
48	2.69 ± 0.09^d	2.13 ± 0.10^d

Normal control value = 1.95 ± 0.06.
Liv.52 control value = 2.20 ± 0.09.
Normal compared with irradiated group column 1 symbols on right side.
The difference between normal control and Liv.52 control was non-significant.
Liv.52 alone compared with Liv.52 + irradiated group column 2 symbols on left side.
Irradiated groups compared with Liv.52 + irradiated groups column 2 symbols on right side.
$^a p < 0.0001$, $^b p < 0.0002$, $^c p < 0.0003$, $^d p < 0.001$, $^e p < 0.002$, $^f p < 0.01$.

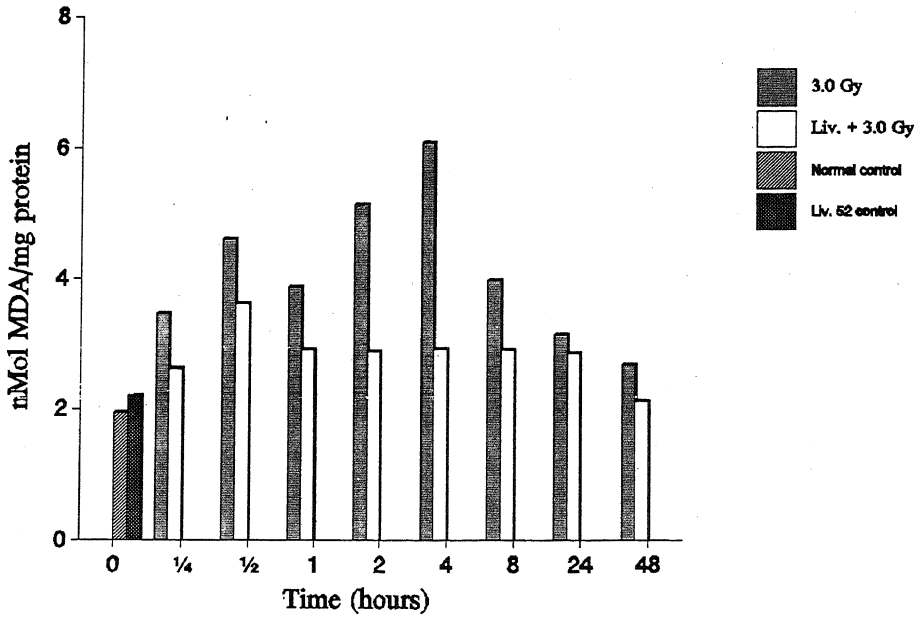


Figure 1. The lipid peroxidation level at different post-irradiation time periods in mouse liver treated or not with Liv.52 before exposure to 3.0 Gy gamma radiation.

Several investigators reported that lipid peroxidation in liver homogenate will start as soon as the supply of endogenous GSH is exhausted, and that the addition of GSH promptly stops further peroxidation[30,31]. It has been reported that the addition of GSH markedly inhibits lipid peroxidation in isolated rat liver microsomes[32].

The administration of Liv.52 before irradiation inhibited the lipid peroxidation significantly as compared to the irradiated control. These results are in accordance with the findings of other investigators, who reported a significant decrease in lipid peroxidation, followed by an increase in the tocopherol level in Liv.52-treated group compared to the CCl_4-treated group[15,16]. Similarly, calmodulin antagonists have been found to inhibit the radiation-induced lipid peroxidation[28]. A depletion in lipid peroxidation was observed in mice erythrocytes treated with MPG before irradiation[29]. WR-1065 and GSH were found to be effective inhibitors of mitochondrial lipid peroxidation induced by HDP/Fe/NADH or by ADP/Fe/ascorbate[33].

Saxena *et al.*[15] reported an increase in the α-tocopherol level in Liv.52-treated rats. In the present study the radiation-induced lipid peroxidation was inhibited significantly by Liv.52 treatment, and this may be due to the elevation of α-tocopherol.

Vitamin E (α-tocopherol) and glutathione are involved in the termination mechanism of lipid peroxidation. Vitamin E donates the hydrogen atom from the chromanol ring hydroxyl group to the lipid radical, generating a tocoperoxy radical. The rate of reaction of vitamin E with lipid hydroperoxy radicals is much faster (~ 10^4) than the rate of reaction of lipid hydroperoxy radicals with neighbouring polyunsaturated fatty acids, thus preventing further damage. Therefore, vitamin E acts as an effective chain breaker, curtailing progression of the lipid peroxidation. However, vitamin E is very low for serving as a lipid radical scavenger, and its concentration in the membrane would rapidly become vanishingly small to reduce the tocoperoxy radical back to tocopherol. Liv.52 may provide α-tocopherol to reduce the tocoperoxy radicals to tocopherol, thereby inhibiting the formation of lipid peroxidation.

Therefore, it is reasonable to assume that radioprotective activity of Liv.52 may be due to the inhibition of lipid peroxidation by increasing the level of α-tocopherol and glutathione.

1. Noda, Y., Ogura, M., Watanabe, S. and Mori, A., *Med. Sci. Res.*, 1993, **21**, 61–62.
2. Tsen, C. C. and Collier, H. B., *Can. J. Biochem. Physiol.*, 1960, **38**, 957–964.
3. Purohit, S. C., Bisby, R. H. and Cundali, R. B., *Int. J. Radiat. Biol.*, 1980, **38**, 147–158.
4. Leyko, W. and Bartosz, G., *Int. J. Radiat. Biol.*, 1986, **49**, 743–770.
5. Petkau, A. and Chelack, W. S., *Biochem. Biophys. Acta*, 1976, **443**, 445–456.
6. Patterson, L. K. and Redpath, J. L., in *Solubilization and Microemulsions* (ed. Mittal, K. L.), Plenum Press, New York, 1997, vol. II, pp. 589–601.
7. Raleigh, J. A., Kremers, W. and Gaboury, B., *Int. J. Radiat. Biol.*, 1977, **31**, 203–213.
8. Helzer, Z., Jozwiak, Z. and Leyko, W., *Experientia*, 1980, **36**, 521–522.
9. Mathur, P. S., *Curr. Med. Pract.*, 1957, **1**, 107–109.

10. Sule, C. R., Pai, V. R., Damania., R. F., *et al.*, *J. Indian Med. Prof.*, 1968, **14**, 6391–6397.
11. Deshpande, R. S., Sheth, S. C. and Joy Kutty, M. D., *Curr. Med. Pract.*, 1971, **15**, 810–816.
12. Saini, M. R., Kumar, S., Jagetia, G. C. and Saini, N., *Indian Pract.*, 1984, **37**, 1133–1138.
13. Jagetia, G. C. and Ganapathi, N. G., *Mutat. Res.*, 1989, **224**, 507–510.
14. Jagetia, G. C. and Ganapathi, N. G., *Mutat. Res.*, 1991, **253**, 123–126.
15. Saxena, A., Sharma, S. K. and Garg, N. K., *Indian J. Exp. Biol.*, 1980, **18**, 1330–1331.
16. Saxena, A. and Garg, N. K., *Indian J. Exp. Biol.*, 1981, **19**, 859–862.
17. Konings, A. W. T. and Drijver, E. B., *Radiat. Res.*, 1979, **80**, 494–501.
18. Bradford, M. M., *Anal. Biochem.*, 1976, **72**, 248–254.
19. Wills, E. D. and Wilkinson, A. E., *Int. J. Radiat. Biol.*, 1970, **17**, 229–236.
20. Yukawa, O. and Nakazawa, T., *Int. J. Radiat. Biol.*, 1980, **37**, 621–631.
21. Bindoli, A., *Free Radiat. Biol. Med.*, 1988, **5**, 247–261.
22. Kerognou, J. F., Braquet, M. and Rcoquet. G., *Radiat. Res.*, 1981, **88**, 377–384.
23. Yukawa, O., Miyahara, M., Shiraishi, N. and Nakazawa, T., *Int. J. Radiat. Biol.*, 1985, **48**, 107–115.
24. Hoffstein, P. E., Hunter, F. E., Gebicki, J. M., Jr and Weinstein, J., *Biochem. Biophys. Res. Commun.*, 1962, **7**, 276–280.
25. Bhanumathi, P. and Uma Devi, P., Proceedings of the 8th International Conference on Chemical Modifiers of Cancer Treatments, Kyoto, Japan, 1993, pp. 123–124.
26. Pompella, A., Romani, A., Fulceri, A. R. E. Maellaro, Benedetti, A. and Comporti, M., in *Eicosanoids, Lipid Peroxidation and Cancer* (eds Nigam, S. K., McBrien, D. C. H. and Slater, T. F.), Springer, Berlin, 1988, pp. 253–263.
27. Maellaro, E., Casini, A. F., Del Bello, B. and Comporti, M., *Biochem. Pharmacol.*, 1990, **39**, 1513–1521.
28. Varshney, R. and Kale, R. K., *Int. J. Radiat. Biol.*, 1990, **58**, 733–743.
29. Ayene, S. I., Kale, R. K. and Srivastava, P. N., *Int. J. Radiat. Biol.*, 1988, **53**, 629–639.
30. Christophersen, B. D., *Biochem. J.*, 1966, **100**, 95–101.
31. Christophersen, B. D., *Biochem. J.*, 1968, **106**, 515–522.
32. Reddy, C. C., Scholz, R. W., Thomas, C. and Massaro, E. J., *Life Sci.*, 1982, **31**, 571–576.
33. Tretter, L., Ronai, E., Szabados, Gy., Hermann, R. ando, A. and Horvath, I., *Int. J. Radiat. Biol.*, 1990, **57**, 467–478.

ACKNOWLEDGEMENTS. We are thankful to Prof. K. Koteshwar Rao, Head, Department of Radiotherapy and Oncology, Kasturba Medical College, Manipal, for providing the necessary irradiation facilities during the course of this study. Financial assistance in the form of senior fellowship from ICMR, New Delhi, to NGG is thankfully acknowledged.

Received 17 July 1995; revised accepted 2 September 1995

An onshore gas well blow-out and its impact observations using satellite data

K. V. S. Badarinath, A. T. Jeyaseelan*, C. S. Jha, Biswadip Gharai and C. B. S. Dutt

Forestry & Ecology Division, *Water Resources Group, National Remote Sensing Agency, Balanagar, Hyderabad 500 037, India

The blow-out of the ONGC Gas Well-19 located at Pasarlapudi in East Godavari District, Andhra Pradesh, attracted lot of concern on its impacts. The blow-out site has been monitored by satellite-based data sets from Indian Remote Sensing Satellite (IRS-P2) and Landsat-5 for impact assessment studies. The satellite data sets showed signature variation in 200 m surrounding the well site. The water umbrella formed over the blow-out site and the resultant temperature and humidity variation measured on the ground were significant up to 200 m radius. Satellite-based optical data have been used for estimating the well temperature. The gas well temperature estimated from satellite data has been found to be around 1100°C. The methodology adopted for estimating the gas well temperature and the results on impact assessment studies are discussed.

THE blow-out of ONGC onshore Gas Well-19 located at Pasarlapudi near Devarlanka in Amalapuram Taluk of East Godavari District, Andhra Pradesh, occurred on 8 January 1995. At about 6.50 PM on that day, there was a sudden increase of gas pressure and the casting was pushed out, with the result that the well caught fire. Initially for about 30 days only the vertical spread of the flame was noticed and subsequently due to damage in the blast of preventor (BOP), which is a vertical structure containing equipment for closing the well in case of exigency, the fire in horizontal direction also increased. Satellite data sets over the region were continuously monitored to obtain cloud-free data over the region and first cloud-free coverage was observed in Indian Remote Sensing Satellite (IRS-P2) pass of 27 January 1995. Spaceborne sensors due to synoptic and repetitive coverage provide data over large regions and have the advantage of providing data over inaccessible regions. The sensors onboard natural resources monitoring missions operate mostly in visible and near-infrared regions of the electromagnetic spectrum. The Landsat satellite data in India are acquired with thematic mapper (TM) sensor having a band in thermal IR region of the electromagnetic spectrum. The thermal band data can be used to study temperature variations below 60°C (ref. 5). In the present study, optical data from IRS-P2 and Landsat TM have been analysed with a view to study the impacts due to blow-out, and ground-based measurements have been collated with the satellite-based obser-

vations. IRS-P2 optical data has been analysed with a view to estimate the well temperature and the methodology adopted has been discussed in detail.

IRS-P2 Linear Imaging Self-Scanning System (LISS-II) digital data of 27 January 1995 and Landsat TM data of 2 February 1995 over the blow-out region have been obtained. Biophysical parameters like leaf temperature, photosynthetically active radiation (PAR), leaf level transpiration and intercellular CO_2 were measured using a leaf chamber analyser (LCA-3); temperature and humidity at different points from the well site were also measured. The vertical profiles of temperature and humidity were measured, respectively, with mercurial thermometer having a sensitivity of 0.1°C in the 0–100°C range and with hygrometer using a stack.

The IRS-P2 LISS-II sensor operates in four bands at 0.45–0.52 μm, 0.53–0.59 μm, 0.63–0.69 μm, 0.77–0.86 μm and Landsat TM operates in seven spectral bands (six in 0.45–2.35 μm region and one in 10.4–12.5 μm thermal IR region).

Satellite data sets recorded in various bands can be analysed for vegetation vigour by ratioing the data of infrared and red bands. The ratio is normally called normalized difference vegetation index (NDVI) and is given by the formula NDVI = (IR-Red)/(IR+Red) (refs 1,4).

The thermal IR data of Landsat are used mainly for studying temperatures below 60°C. In the present study an alternate method for estimating the temperature of the gas well has been adopted[5]. The short wavelength region of 0.4–3.0 μm is not commonly used for monitoring the normal thermal regime of the earth's surface as it emits a very small quantity of radiation in this spectral region. In the case of gas well burning, the observed high-temperature region acts like a source of radiation and the radiance values recorded by satellite-based sensors correspond to the emitted radiation. This concept enables the use of Planck's function for estimating the temperature over the gas well. Similar studies have been reported in the literature over volcanic regions[2,3,5].

The false colour composite of the study area generated from TM bands 4, 3, 2 (RGB) is shown in Figure 1 *a* and that from TM bands 4, 5, 7 is shown in Figure 1 *b*. The enlarged false colour composites corresponding to these are shown in Figures 1 *d* and 1 *e*, respectively. The blow-out region can be clearly seen as a bright yellow spot in Figure 1 *d*, which is an enlarged image of Figure 1 *a*. In Figure 1 *b*, severe stress over the vegetation due to heat from blow-out is indicated by bluish grey area around the well site as the TM bands

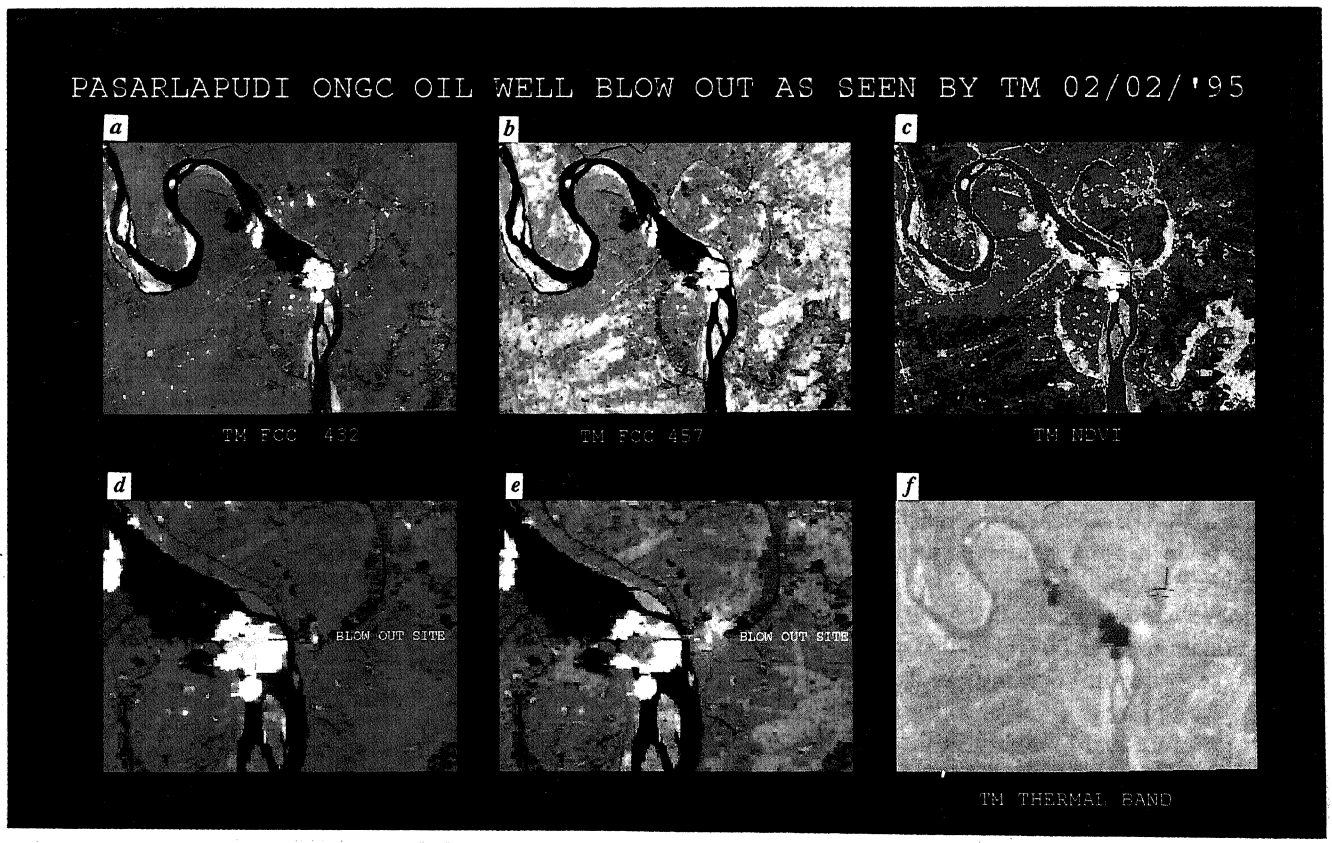


Figure 1. Thematic mapper data of the study area on 2 February 1995: *a*, False colour composite (FCC) of bands 4, 3, 2; *b*, FCC of bands 4, 5, 7; *c*, normalized difference vegetation index (NDVI); *d*, enlarged FCC of bands 4, 3, 2; *e*, enlarged FCC of bands 4, 5, 7; *f*, thermal IR image.

4, 5, 7 (0.77–0.86 μm, 1.55–1.75 μm and 2.08–2.35 μm) are sensitive to vegetation stress. The NDVI image generated from bands 4 and 3 (0.77–0.86 μm and 0.63–0.69 μm) is shown in Figure 1 *c*, wherein the nongreen area around the blow-out site indicates the stressed area. The thermal image of TM band 6 (10.4–12.5 μm) is shown in Figure 1 *f* and it can be observed that due to saturation of thermal radiance values above 60°C, the blow-out site region appears bright and it is difficult to pinpoint the well site.

The satellite observations of localized vegetation stress around the well site were cross-checked with ground observations. Ground measurements of air and soil temperature, humidity, crop photosynthetic rate, transpiration and radiation near the well site were made using the LCA-3 instrument. Further, discussions with Revenue, Agriculture and ONGC officials were made to understand the economic impacts.

The satellite data analysis indicates localized damage of around 200 m close to the well site, as can be seen in Figure 1, and is in conformity with ground observations. The ground observations showed that the temperature was around 40°C at 24–30 f height close to the well site (150 m) and the ground temperature was at 32°C. The temperature was uniform with surroundings beyond 150 m, mainly due to the heat protection by coconut plantations. The well site (BOP) was totally provided with water umbrella by ONGC using jet pumps. The moisture thus generated formed a localized cloud and can be seen as a cloud patch appearing in bright white in Figure 1. The biophysical parameters showed normal values in standing paddy crops close to the blow-out site (150 m) as the ground temperatures were normal and sufficient stagnated water was available in the fields. The coconut trees within 200 m radius showed damage signs as the temperatures were high above ground level due to the gushing flames. The coconut saplings facing BOP at 150 m distance showed yellowing and browning due to damage of leaves. In the leeward side of BOP, healthy coconut saplings are observed. The persistent high temperature of above 40°C at the coconut crown led to falling of fruits near the BOP site (200 m). The heat-tolerant Ficus sp. plant was found to be in healthy state even at a distance of 150 m from the well site.

The observations made by A. P. Pollution Control Board, Rajahmundry, indicated that no poisonous gases were present due to burning in the atmosphere. The physiological process level changes due to exposure to high temperature and humidity in the less impacted area were also shown to be normal. Noise pollution was in the range of 100–140 dB close to the well site and 55 dB (normal) around 2 km radius, as reported by A. P. Pollution Control Board.

The temperature of the well site was estimated to be around 1100°C; this value is close to the reported values. The effect of emitted gases and water vapour on the satellite-derived radiance values has not been considered while calculating the temperature. The ground temperature and humidity variation at different heights from the well site is shown in Figure 2. It can be observed that the temperature values are increasing with height, due to the gushing flames spreading above ground level. Due to water umbrella, the humidity values were found to be high in the windward direction away from the well site.

The analysis of satellite data suggested localized effects (around 200 m regions) due to the oil well blow-out. The temperature values estimated from the optical data are in conformity with ONGC ground observations, suggesting the possible use of optical data for studying high-temperature phenomena. The noise pollution levels are high within 2 km region around the site. The damage to coconut plantations in down wind direction were

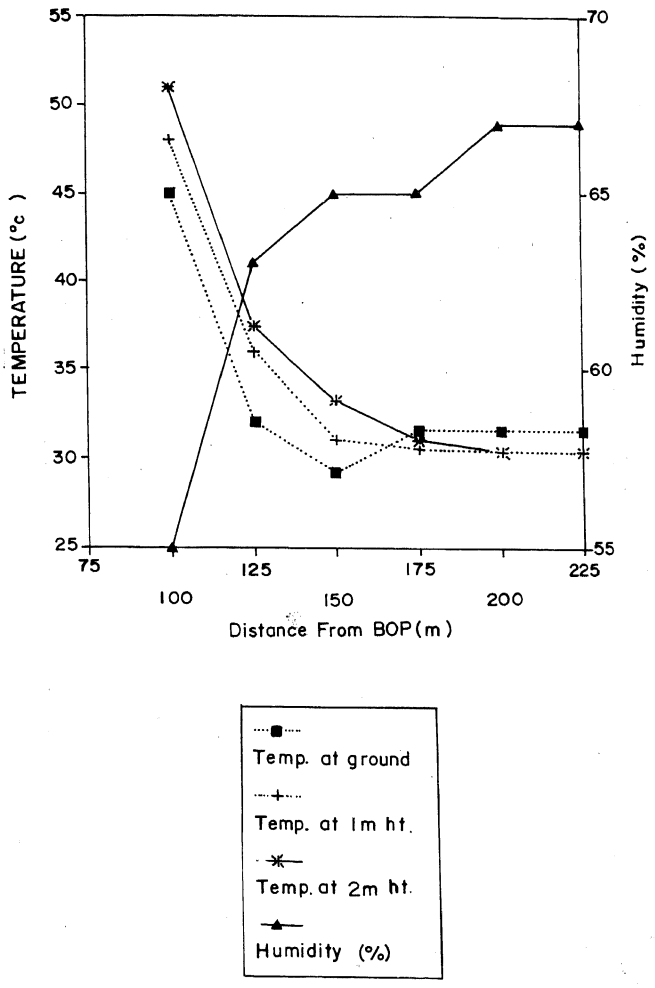


Figure 2. Temperature and humidity profiles at various points from the blow-out site.

observed within 200 m zone. The area under 200 m radius from BOP is the zone of influence of blow-out. The gas blow-out site was inaccessible and satellite data alone facilitated in observing the vent temperature synoptically during the blow-out condition. In addition, multitemporal satellite data by virtue of repetitive coverage facilitated dynamic aspects of blow-out condition and the associated spatial changes resulting due to damage. Thus, satellite technology only could facilitate synoptic observation of such episodic events and could account spatially for the area of its influence.

1. Goward, S. N., Markham, B., Dye, D. G., Dulaney, W and Yang, J., *Remote Sens. Environ.*, 1991, **35**, 257–277.
2. Rothery, D. A., Francis, P. W. and Wood, P. A., *J. Geophys. Res.*, 1988, **93**, 7993–8008.
3. Rothery, D. A., *Int. J. Remote Sens.*, 1989, **10**, 1423–1427.
4. Justice, C. O., Townshend, J. R. G., Holben, B. N. and Tucker, C. J., *Int. J. Remote Sens.*, 1985, **6**, 1271–1318.
5. Gupta, R. K. and Badarinath, K. V. S., *Int. J. Remote Sens.*, 1993, **14**, 2907–2918.

ACKNOWLEDGEMENTS. We express our sincere gratitude to Prof. B. L. Deekshatulu, Director, NRSA, and Dr D. P. Rao, Associate Director, NRSA, for their encouragement. We are also grateful to Dr S. T. Chari, Group Director, Water Resources Group, NRSA, for the technical discussions.

Received 17 May 1995; revised accepted 28 August 1995

Oxygen linkages in neodymium heptamolybdate single crystals – An XPS study

Sushma Bhat, P. N. Kotru and K. S. Raju*

Department of Physics, Jammu University, Jammu 180 001, India
*Department of Crystallography and Biophysics, University of Madras, Guindy Campus, Madras 600 025, India

Single crystals of neodymium heptamolybdate ($Nd_2Mo_7O_{24} \cdot 27H_2O$) are grown in sodium metasilicate gel by the diffusion of neodymium nitrate (upper reactant) into the set gel impregnated with a mixture of molybdenum trioxide, ammonium hydroxide and concentrated nitric acid (lower reactant). Oscillation X-ray diffraction pattern shows that neodymium heptamolybdate is single-crystalline in nature, while infrared absorption spectrum establishes the presence of molybdate group and water of crystallization. X-ray photoelectron spectroscopic studies of neodymium heptamolybdate single crystals establish the presence of neodymium and molybdenum in their oxide states and the oxygens in the sample are shown to exist as (i) terminal oxygen (Mo=O), (ii) bridging oxygen (Mo–O–Nd) and (iii) oxygen of lattice water (H–O–H).

FERROELECTRIC and ferroelastic properties of rare-earth molybdates are of great interest in electro-optical and acousto-optical devices[1–3]. Barium molybdate tetragonal bipyramidal crystals are grown by the precipitation of alkaline-earth metal molybdate powders from neutral aqueous solutions[4]. Rare-earth molybdates, having the general formula $R_2(MoO_4)_3$, have been reported to be ferroelectric materials[5]. The growth of rare-earth molybdates (R = Pr, Nd, Sm, Eu, Tb and Dy) using the Czochralski technique has been studied by Brixner[6]. As this method involves elevated temperatures, thermal stresses released during the growth may make the crystals defective. Henisch and coworkers[7,8] have established the growth of crystals in gels and this method has been fully exploited in the case of calcium sulphate dihydrate single crystals[9–15]. Growth and characterization of rare-earth mixed single crystals of samarium barium molybdate have been reported by Isac and Ittyachen[16]. Studies on the growth of neodymium heptamolybdate (hereafter called NHM) single crystals in silica gels have been reported in the literature[17].

For the growth of NHM in gels in the test tube, acidified aqueous solution of sodium metasilicate (0.5 M, pH 5) is used as the reacting medium. The lower reactant (to be incorporated inside the gel before setting) is prepared as follows: 0.5 M molybdenum trioxide is completely dissolved in 15 N ammonium hydroxide using a magnetic stirrer, to which nitric acid (24 ml in 50 ml distilled water) is added. This lower reactant is mixed with the above gel and allowed to set undisturbed (gel age 96 h). On top of this set gel, as the upper reactant, a solution of 0.75 M neodymium nitrate is slowly added through the wall of the test tube, so as not to rupture the surface of the set gel. In about three weeks' time, platelets of NHM (of sizes 2 mm × 2 mm × 1 mm) appear below the thick white layer of the precipitate (of NHM) inside the gel.

The mechanism of chemical reactions leading to the formation of NHM is explained as follows:

$$MoO_3 + 2NH_4OH \rightarrow (NH_4)_2 MoO_4 + H_2O. \quad (1)$$

When acidified, we get

$$(MoO_4)^{2-} + H^+ \rightarrow HMoO_4^- , \quad (2)$$

$$HMoO_4^- + H^+ + 2H_2O \rightarrow Mo(OH)_6 , \quad (3)$$

$$Mo(OH)_6 + HMoO_4^- \rightarrow (HO)_5 MoOMoO_3 + H_2O, \quad (4)$$

$$(HO)_5 MoOMoO_3^- + HMoO_4^- \rightarrow (HO)_4 Mo(OMoO_3)_2^{2-} + H_2O, \quad (5)$$

*For correspondence.

$$(HO)_4\,Mo(OMoO_3)_2^{2-} + 3HMoO_4^- \rightarrow (HO)\,Mo(OMoO_3)_5^{5-} + 3H_2O, \quad (6)$$

$$(HO)\,Mo(OMoO_3)_5^{5-} + HMoO_4^- \rightarrow Mo(OMoO_3)_6^{6-} + H_2O, \quad (7)$$

$$Mo(OMoO_3)_6^{6-} \rightarrow Mo_7O_{24}, \quad (8)$$

$$2Nd^{3+} + Mo_7O_{24}^{6-} \rightarrow Nd_2Mo_7O_{24}. \quad (9)$$

The oscillation X-ray diffraction (XRD) pattern (Figure 1) of NHM platelet was recorded in a Unicam camera with Cu K_α radiation ($\lambda = 1.5418$ Å). The revelation of well-defined Bragg reflections at specific (2θ) angles confirms that NHM is ordered, meaning single crystalline, in nature.

The IR absorption spectrum of NHM taken in KBr matrix confirms the presence of the water of crystallization by the well-defined broad peak at 3000–3600 cm^{-1} (relating to antisymmetric and symmetric OH sretchings) and a peak at 1630 cm^{-1} (relating to HOH bending). The presence of molybdate groups in NHM is also established by the revelation of a pronounced broad peak at 810 cm^{-1} relating to symmetric or antisymmetric stretching modes of vibration[18] and a peak at 325 cm^{-1} relating to symmetric bending modes of vibration[19].

It is worth mentioning, that in the present work, so far, neither XRD nor infrared (IR) absorption spectrum has been able to identify heavy elements like Mo and Nd unambiguously. Hence, it was thought worthwhile to take up the most sophisticated X-ray photoelectron spectroscopic (XPS) technique.

In XPS (more commonly known as ESCA), the identification of the elements incorporated in the sample is established from the binding energy (BE) values (eV) relating to the peaks of their respective elements, while the quantitative data are obtained from the peak heights or areas. The chemical states of the elements were identified from the exact positions and separations of the peaks, as well as from certain spectral contours. XPS studies were carried out on NHM single crystals in order to confirm the presence of heavy elements (Nd and Mo in the present case) besides their chemical states, which, consequently, sheds information on the linkages of the oxygens as well.

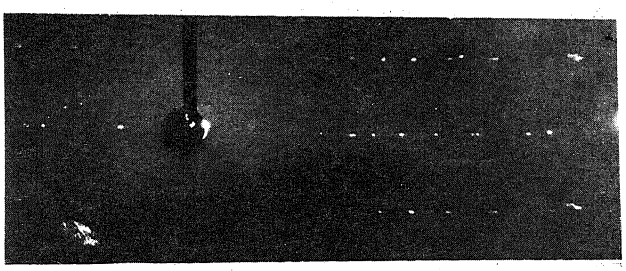


Figure 1. Oscillation XRD pattern of NHM.

The XPS investigations of NHM were carried out at the Regional Sophisticated Instrumentation Centre, Indian Institute of Technology, Powai, Bombay, using a system VG scientific ESCALAB MK2 spectrometer (Al K_α source, photon energy 1486.6 eV) operated at 10 kV and 10 mA. The overall resolution of the system with an analyser pass of 50 eV is about 0.7 eV. The pressure in the analyser chamber was maintained below 10^{-9} torr.

In the present investigations, the BE (eV) scale (X-axis) was calibrated for C 1s (284.6 eV), which is due to the contribution of adventitious hydrocarbon from the laboratory environment or from the glove box.

The XPS peaks relating to Mo 3d states (Figure 2 *a*) are positioned at BEs of around 232.5 and 235.7 eV, with the separation of 3.2 eV strongly suggesting that Mo in NHM is in the oxide state (MoO_3). Similarly, the XPS peak relating to $3d_{5/2}$ state of Nd (Figure 2 *b*) in NHM has a BE of around 980 eV. This BE value confirms that Nd too is in its oxide state (Nd_2O_3) in the NHM sample, which is in good agreement with literature findings[20] on the structural characterization of

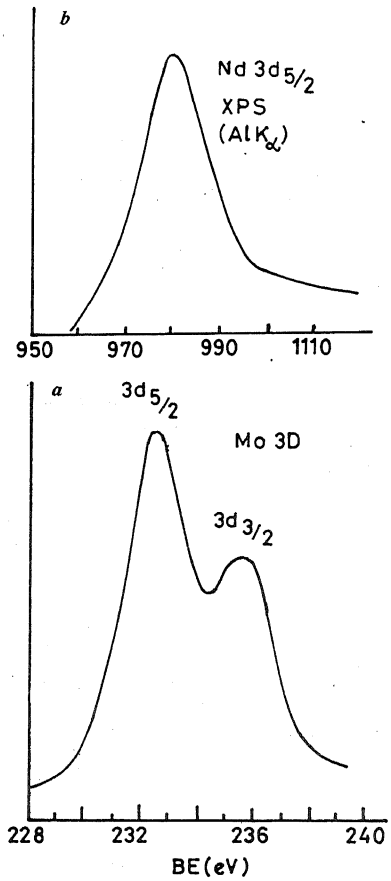


Figure 2. XPS peaks of (*a*) 3d levels of molybdenum and (*b*) $3d_{5/2}$ level of neodymium.

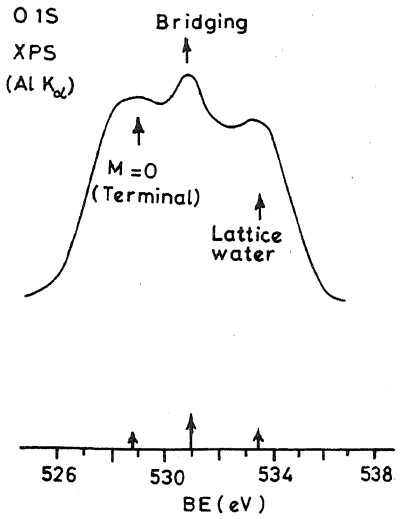


Figure 3. XPS peak of oxygen 1s.

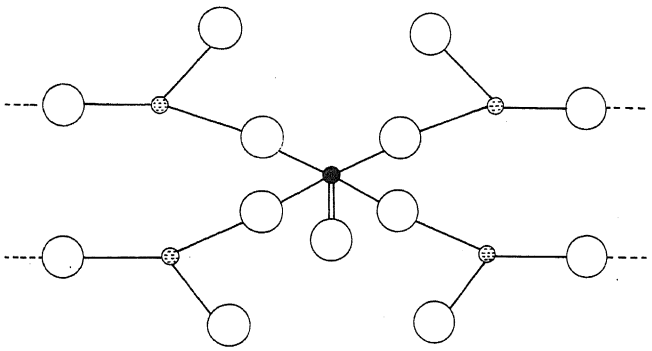


Figure 4. Schematic of oxygen coordinations in NHM: molybdenum (small dark circle), neodymium (small dashed circles) and oxygen (big blank circles).

gel-grown neodymium copper oxalate single crystals.

It is interesting to note that the XPS peak relating to O 1s (Figure 3) has a complex spectral contour and is a broad peak comprising of three peaks. These peaks relate to BEs around 528.8, 531 and 533.5 eV, respectively, giving the clue that oxygens in NHM have three different coordinations.

The incorporation of Mo, Nd and O in NHM has been confirmed by XPS (Figures 2 and 3) besides establishing the existence of Mo and Nd as their oxides (Figure 2 *a*, *b*). These conclusions are strongly supported by the binding energy values corresponding to XPS peaks of Mo(3d), Nd($3d_{5/2}$) and O(1s) (Figures 2 and 3), which are in excellent agreement with literature findings[21].

Attention is drawn to the fact that the chemical formula for NHM ($Nd_2Mo_7O_{24}$) as arrived at from the mechanism of reaction kinetics (eqs (1)–(9)) seems reasonable, as to be having the entities (Nd_2O_3) and 7(MoO_3), as supported by the XPS peaks of Nd and Mo in their oxide states (Figure 2). The possible oxygen linkages with MoO_3 and Nd_2O_3 in NHM are shown schematically in Figure 4. A careful examination of Figure 4 reveals that oxygens have two types of linkages, namely, terminal oxygen (Mo=O) and bridging oxygens (Mo–O–Nd). In order to account for the third entity, i.e. water of crystallization, the third peak in the complex spectral contour of XPS of O 1s (Figure 3) is attributed to the oxygen getting coordinated to the hydrogens of lattice water (H–O–H), whose presence is confirmed by the IR absorption spectrum of NHM. Thus, XPS studies throw light on the presence of Mo, Nd and O as well as on their chemical states, confirming the speculation of the entities of NHM as MoO_3, Nd_2O_3 and lattice water. Besides this, complex spectral contour of XPS peak of O 1s has been helpful in understanding the linkages of oxygens in NHM as well (Figure 4).

The heights and/or areas of XPS peaks, meaning their intensities, are measures of the quantities of the species present in the material[22]. It is clearly evident from Figure 3 that the intensity of O 1s relating to bridging oxygens is higher than that of terminal oxygens. Hence, it may be concluded that, from the viewpoint of structural aspects, the contribution of the bridging oxygens is more significant than that of the terminal oxygen in NHM.

1. Barkley, J. R., Brixner, L. H. and Horgan, E. M., IEEE Symp. on the Applications of Ferroelectrics, York Town Heights, New York, 1971.
2. Barkley, J. R., Brixner, L. H., Horgan, E. M. and Waring, R. K., *J. Ferroelec.*, 1972, **3**, 191–195.
3. Sapriel, J. and Vacher, R., *J. Appl. Phys.*, 1977, **48**, 1191–1194.
4. Pacter, A., *Krist. Tech.*, 1997, **12**, 729–732.
5. Borchordt, H. J. and Bierstedt, F. E., *Appl. Phys. Lett.*, 1966, **8**, 50–54.
6. Brixner, L. H., *J. Cryst. Growth*, 1973, **18**, 297–302.
7. Henisch, H. K., *Crystal Growth in Gels*, Pennsylvania State University Press, University Park, Pennsylvania, 1970.
8. Henisch, H. K., Dennis, J. and Hanoka, J. I., *J. Phys. Chem. Solids*, 1965, **26**, 493–500.
9. Raju, K. S., *Cryst. Res. Technol.*, 1983, **18**, 1277–1281.
10. Jayakumar, D. and Raju, K. S., *Bull. Mater. Sci.*, 1983, **5**, 339–404.
11. Raju, K. S., *J. Mater. Sci. Lett.*, 1983, **2**, 705–709.
12. Raju, K. S., *J. Mater. Sci.*, 1981, **16**, 2512–2516.
13. Raju, K. S., *J. Mater. Sci.*, 1985, **20**, 756–760.
14. Raju, K. S. and Pippel, E., *Phys. Stat. Solidi A*, 1984, **81**, K93–K95.
15. Raju, K. S., Godehardt, R., Hoppfe, J. and Pipple, E., *Cryst. Res. Technol.*, 1984, **19**, 1127–1131.
16. Isac, Jayakumari and Ittyachen, M. A., *Bull. Mater. Sci.*, 1992, **15**, 349–353.
17. Bhat, Sushma, Kotru, P. N. and Koul, M. L., *Mater. Sci. Engg. B*, 1993, in press.
18. Busy, R. H. and Keller, O. L., Jr, *J. Chem. Phys.*, 1964, **41**, 215–225.
19. Clark, G. M. and Doyle, W. P., *Spectrochim. Acta*, 1966, **22**, 1441–1447.
20. Raju, K. S., Krishna, K. N., Isac, Jayakumari and Ittyachen, M. A.,

Bull. Mater. Sci., 1994, **17**, 1447–1455.
21. Wagner, C. D., Riggs, W. M., Davis, L. E., Moulder, J. F. and Mullenberg, J. E. (eds), *Handbook of X-ray Photoelectron Spectroscopy*, Perkin Elmer, Physical Electronics Division, Minnesota, 1978.
22. Riggs, W. M. and Parker, M. J., in *Methods of Surface Analysis* (ed. Czanderna, A. W.), Elsevier, Amsterdam, 1975, Chap. 4, p. 108.

ACKNOWLEDGEMENTS. We thank the Regional Sophisticated Instrumentation Centre, Indian Institute of Technology, Powai, Bombay, for the ESCA facility. KSR thanks Prof. B. Viswanathan, Department of Chemistry, Indian Institute of Technology, Madras, for stimulating discussions.

Received 27 March 1995; revised accepted 5 June 1995

Tectonic evolution of the Central Gujarat plain, western India

D. M. Maurya, L. S. Chamyal and S. S. Merh

Department of Geology, Faculty of Science, M.S. University of Baroda, Baroda 390 002, India

The Quaternary basin of the Gujarat plain is formed due to the reactivation of Tertiary basement faults. The uplift as a horst of the Aravalli in the east and the Saurashtra peninsula in the west, comprising a culminating event of the post-Mesozoic rifting of the western continental margin, resulted in the Quaternary basin which was later filled up by alluvial sediments. The basin that received the fluvial sediments marked the end phase of the Cambay graben subsidence that was initiated in the Paleocene. The nature of variation in thickness of the Quaternary sediments in its different segments indicates that the basin comprised a series of horsts and grabens. Whereas the Early Quaternary tectonism gave rise to the differential basement topography, the tectonic events of the Late Quaternary were responsible for the present landscape development in Central Gujarat.

THE 300–800 m thickness of Quaternary fluvial and aeolian sediments resting upon the Tertiary rocks of the Cambay basin is the consequence of reactivation of pre-Quaternary basement faults. The Quaternary basin is delimited by marginal faults in the west and east. To the SE lies the Narmada–Son Lineament, and in the NE is the Aravalli horst (Figure 1). The basin comprises several fault-bound structural blocks[1] which continue up to Moho boundary[2]. The Gujarat Quaternary sediments have been studied by several workers (Allchin *et al.*[3], Chamyal[4], Chamyal and Merh[5], Merh and Chamyal[6] and Pant and Chamyal[7]); their emphasis, however, was mainly on the processes and agents of deposition, and on the paleoclimatic variations recorded in sediments. Recent work has demonstrated that successive tectonic events in the Gujarat region since Cretaceous[8] have also controlled the pattern of sedimentation and nature of the sediment body. The Cambay and Narmada basins, according to Biswas[8] opened up as a result of the counterclockwise rotation accompanying drift of the Indian plate.

The Gujarat plains, in the light of the overall Cambay basin tectonics, point to an important role played by differential movements along numerous structural lineaments in controlling the process of filling up of the Quaternary basin and sculpturing the present-day landscape. The movements along tectonic lineaments at various scales are significantly reflected in (i) the disruption of an ancient superfluvial system and its replacement by the existing drainage[9], (ii) development of ravines along various river courses indicating post-depositional fracturing, (iii) differential altitudes of river terraces, (iv) capturing of rivers (e.g. Rupen by Sabarmati, Orsang by Narmada), (v) differential uplifts along most major rivers giving rise to a distinct northward tilting of the respective left banks, and (vi) progressively increasing subsidence from south to north[10].

We have for the first time attempted to interpret the role of tectonics in the evolution of the Quaternary basin and its subsequent filling. Obviously, the Quaternary sedimentation is found to have been controlled by the same tectonic features which were responsible for the deposition of the Tertiaries. With the withdrawal of the Tertiary sea in this part, marine processes were replaced by fluvial sedimentation, marking the onset of the nonmarine continental depositional cycles. The marine sediments of Upper Tertiary show a transition[11] changing over to a period dominantly of fluvial sedimentation during the Quaternary.

The basin that received Quaternary sediments, obviously, comprised the more or less filled up Tertiary Cambay basin and marked the culminating phase of the Cambay and Narmada graben tectonics, though the tectonism preceding, during and post-dating the continental deposition was an integral part of the Cambay basin tectonism. ONGC investigations have adequately shown that the Cambay basin formed during Paleocene subsided episodically with varying rates and finally during Quaternary the rate of subsidence slowed down. It is however, not unlikely that the Quaternary basin experienced considerable reactivation of the pre-existing faults. These faults were an integral part of the Cambay basin rift phenomenon. It is possible that the advent of Pleistocene witnessed a renewal of tectonic activity and triggering of a new cycle of subsidence. This neotectonism coincided with the uplift of Aravalli in the east and northeast[12]. An uplift of the order of 300 m of the Aravalli has been envisaged by Ahmad[13] during Quaternary. To

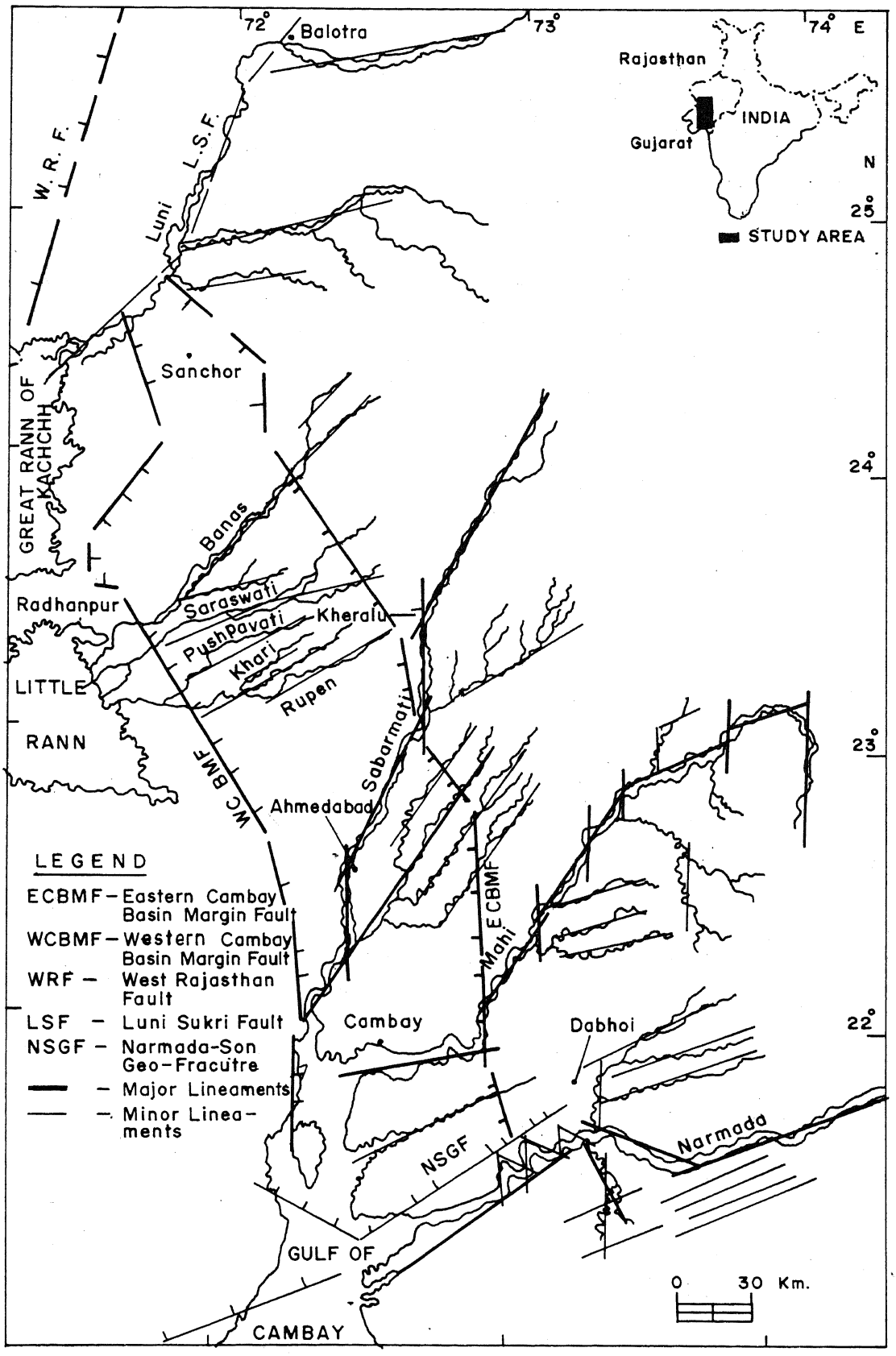


Figure 1. Tectonic map of the Gujarat alluvial plain superimposed over the tectonic map of Cambay basin[14].

the west, the Saurashtra peninsula formed an uplifted horst (Figure 2). This renewal of tectonic activity was essentially related to the Cambay basin faulting and the Narmada geofracture.

The development of a rift-related depression in front of a rising hill range provided an ideal site for the initiation of fluvial sedimentation. The exposed river sections as well as subsurface data provide clear and unequivocal evidences pointing to an enormous thickness varying from about 800 m in the south to 300 m in the north. The variation in thickness is controlled by basinal faults which were effective prior to and during the sedimentation. Significantly, colluvial deposits are absent and fluvial material is encountered right from the rocky fringes. Although the sediment thickness varies from block to block (Figure 3), in composition the fluvial sediments comprise gravels, coarse to fine sands, silts and clays, and point to their deposition in several instalments, depositional process punctuated by periods of nondeposition and pedogenesis. The sediment nature and thickness point to an overall basin subsidence, but one cannot be very sure about the variations in the movement and in the course of Quaternary too, it is not unlikely that inversions in movement directions did take place. Such a tectonic pattern has been established for the trappean highlands of South Gujarat by Alavi and Merh[10], and similar post-Tertiary tectonic activity can safely be visualized.

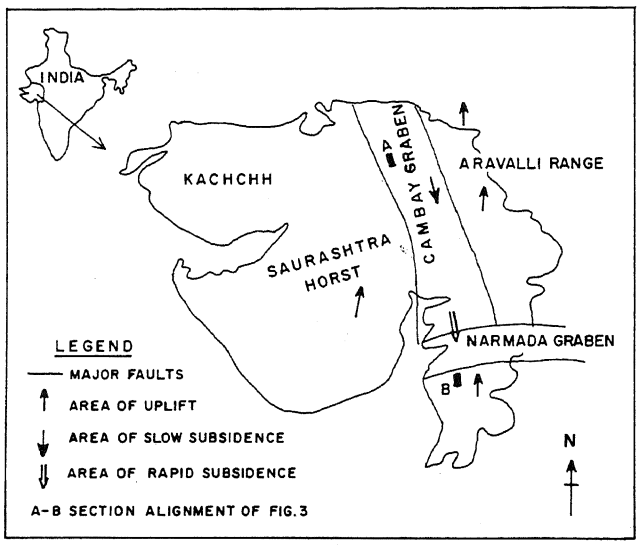


Figure 2. Sketch map of Gujarat showing major tectonic movements during Quaternary.

The cross-section prepared based on ONGC data clearly brings out the influence of the tectonic framework of the Cambay and Narmada basins along with their various tectonic blocks on the Quaternary sediments (Figure 3). The basin therefore, comprised series of horsts and grabens which were the direct manifestations of the various Tertiary structural blocks. These tectonic sub-basins were subjected to periodic uplifts and subsidences relative to each other, while overall the basin was subsiding. Whereas the ONGC workers have worked out in considerable detail the role of tectonics in Tertiary sedimentation, no information was provided by them on the Quaternary history. There is, however, little doubt that the Late Quaternary tectonism has been responsible for the present landscape development in Central Gujarat.

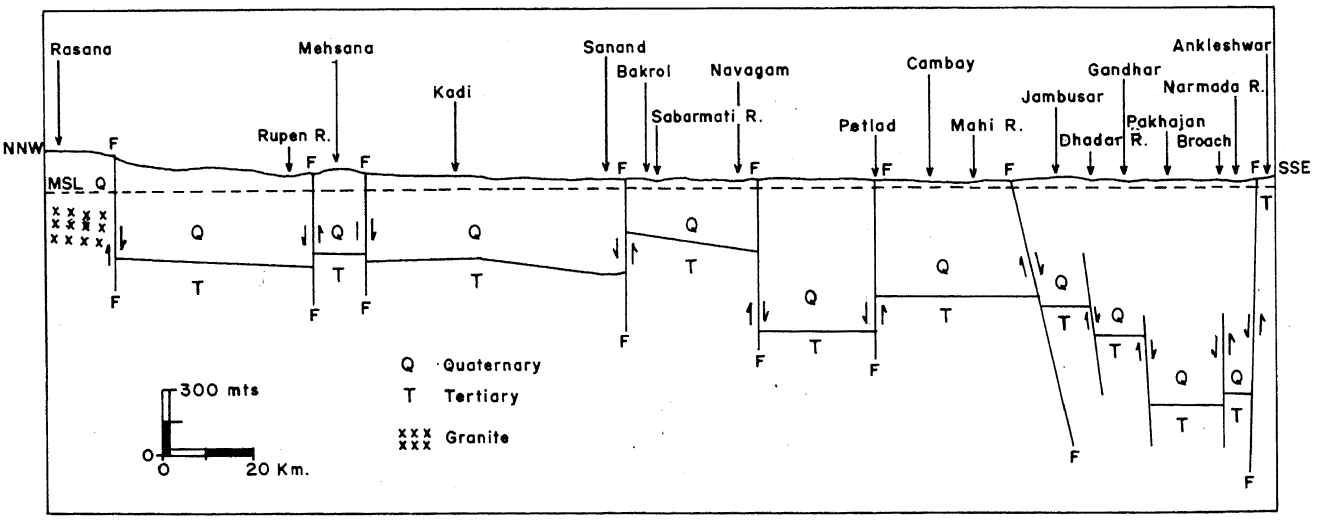


Figure 3. Schematic cross-section of the Quaternary basin (based on thickness data obtained from ONGC archives).

1. Chandra, P. K. and Chowdhary, L. R., *Bull ONGC India*, 1969, **6**, 37–50.
2. Kaila, K. L., Krishna, V. G. and Mali, M. D., *Tectonophysics*, 1980, **76**, 99–130.
3. Allchin, B., Goudie, A. and Hegde, K. T. M., *The Prehistory and Palaeogeography of the Great Indian Desert*, Academic Press, London, 1975.
4. Chamyal, L. S., *J. Geol. Soc. India*, 1995, **45**, 345–348.
5. Chamyal, L. S. and Merh, S. S., *Man Environ.*, 1992, **27**, 33–40.
6. Merh, S. S. and Chamyal, L. S., *Curr. Sci.*, 1993, **64**, 823–827.
7. Pant, R. K. and Chamyal, L. S., *Proc. Indian Natl. Sci. Acad.*, 1990, **56**, 501–511.
8. Biswas, S. K., *Tectonophysics*, 1987, **135**, 315–327.
9. Sridhar, V., Chamyal, L. S. and Merh, S. S., *J. Geol. Soc. India*, 1994, **44**, 427–434.
10. Alavi, S. A. and Merh, S. S., *Proc. Indian Natl. Sci. Acad.*, 1991, **57**, 683–698.
11. Merh, S. S., *J. Geol. Soc. India*, 1993, **41**, 259–276.
12. Sen, D. and Sen, S., *Tectonophysics*, 1983, **93**, 75–98.
13. Ahmad, F., *Proc. Indian Natl. Sci. Acad.*, 1986, **52**, 1286–1306.
14. Markovich, V. P., unpublished ONGC Report, 1976.

ACKNOWLEDGEMENTS. We are indebted to Prof. K. S. Valdiya for critically going through the manuscript. Financial assistance provided by Department of Science and Technology is gratefully acknowledged.

Received 25 July 1994, revised accepted 14 August 1995

Occurrence and petrogenesis of Loda Pahar trondhjemitic gneiss from Bundelkhand craton, central India: Remnant of an early crust

Kewal K. Sharma and Abdul Rahman

Wadia Institute of Himalayan Geology, Dehra Dun 248 001, India

The occurrence and the geochemistry of the typical high-Al trondhjemitic gneiss intruding into the highly deformed basic host rocks from Loda Pahar, Kabrai area, northeastern part of the Bundelkhand massif is reported. This large relict of the early crust is intruded by various phases of Proterozoic calc-alkaline granitoids. The highly fractionated REE pattern with HREE depletion and concave upward shape at the HREE end of the Loda Pahar trondhjemitic gneiss (LPTG), besides the major and trace element chemistry, is typical of Archaean TTG suites reported from shield areas of other continents. The evolution of such trondhjemitic magma is considered to have played a significant role during early crustal growth and the LPTG provides the first convincing evidence to qualify the Bundelkhand craton as a typical Archaean terrain.

IN recent years considerable work has been done to understand the nature and evolutionary trends in the formation of early crustal rocks, the relicts of which are exposed in shield areas of all the continents. The Archaean crust in general is comprised of granitoid gneisses, massive granitoids and greenstones in proportion of 6 : 3 : 1[1]. The evolutionary trends recognized in the granitoid rocks of the Precambrian terrains of the world contain two main associations, an older Na-rich tonalite–trondhjemite–granodiorite (TTG) suite and younger K-rich granodiorite–monzonite–adamellite–granite (GMAG) suite. The former predominates throughout the Archaean, but subsequently, the latter becomes significant. The chemical and isotopic characters of the two suites are distinct and provide a major temporal demarcation in continental evolution at the Archaean–Proterozoic boundary[1].

In India, evidence to reconstruct different stages of early crustal evolution have so far come from the cratonic nucleii of Dharwar, Singhbhum and very recently from Bastar. Though the Bundelkhand craton is being considered unique for representing an early crust, unaffected by Proterozoic and subsequent deformational events[2], there is no convincing geochemical or geochronological data to indicate the presence of an early crust from Bundelkhand craton except for the geochemical data of early Proterozoic Bundelkhand granitoids[3].

The authors initiated detailed petrochemical studies of the Bundelkhand craton and have identified different components (metabasic and metasediments with banded iron intruded by various phases of TTG) of the early crust into which the Proterozoic calc-alkaline granitoid suite intrudes. The present paper reports the occurrence of a large relict block of trondhjemitic gneiss in deformed Archaean crust intruded by undeformed early Proterozoic calc-alkaline granitoids from Loda Pahar, Kabrai area. It also discusses the geochemistry, petrogenesis and its implication to the evolution of the Archaean crust in the Bundelkhand craton of the Central Indian Shield.

Semicircular to triangular outcrop of Bundelkhand granite massif in the northern part of the Indian Shield is known to be comprised dominantly of granitoid plutons of various phases, younger basic dykes and quartz reefs, mainly the early Proterozoic magmatic events. The information on the nature of the Archaean crust into which the younger Proterozoic granitoids intruded is very much limited to the reported occurrences of numerous enclaves of schists, gneisses, banded iron, calc-silicates, mafic and ultramafic rock suites only[4].

Recently, Rahman and Zainuddin[3] identified five phases within the Bundelkhand massif and provided geochemical data to characterize their source rock from Mahoba area. The earliest phase of the hornblende-bearing granodiorite having metaluminous composition and Y depletion has been attributed to hydrous partial melting of hornblende- and/or garnet-bearing mafic source. The

porphyritic varieties of the biotite granite which intrude into the hornblende granodiorite are Y-undepleted, I-type, monzo-granites and were generated from the melting of the felsic crust. The subsequent intrusion of two phases of younger leucogranites (coarse and fine-grained varieties) are peraluminous monzo- to syno-granite and show marked enrichment with Y, Zr, Th, K and Rb and depletion of P, Sr, Ti and Ba. The magma for these leucogranite phases corresponds to the highly fractionated calc-alkaline to alkaline composition and felsic I-type granite characteristics. Following the preliminary geochronological age of 2550 Ma (based on Rb–Sr whole-rock isochron of seven samples collected from widely separated areas from the massif) for the Bundelkhand granite by Crawford[5], Sarkar *et al.*[6] carried out detailed Rb–Sr whole-rock dating of various plutons exposed in the northwestern part (Jhansi–Babina track). Six whole-rock samples of grey granodiorite–granite suite from a quarry define a Rb–Sr isochron age of 2359 ± 53 Ma with initial Sr ratio Sr_i of 0.7082 ± 0.0036. Another set of five whole-rock samples of pink-coloured coarse-grained granite–granite porphyry–aplite suite from Babina–Talbehat yielded a Rb–Sr isochron age of 2246 ± 78 Ma with Sr_i ratio of 0.7125 ± 0.0045. Four whole-rock samples of Badera gneissic migmatites (southeast of Babina) yielded a Rb–Sr isochron age of 2130 ± 102 Ma, possibly reflecting a strong thermal imprint due to large-scale emplacement of younger (pink) granite phase rather than the true age of the formation of the gneissic component of the older crust. However, the Geological Survey of India has dated 3.5-Ga-old gneissic rock from Babina–Talbehat area[7].

The authors during their initial field studies recognized numerous xenoliths of highly folded gneissic rocks within

Table 1. Geochemical composition of trondhjemitic gneiss from Bundelkhand Craton, Central Indian shield

	G-21	G-22	G-23	G-24	G-25	G-26	G-27
SiO_2	70.69	70.01	70.39	69.94	69.45	70.12	69.96
TiO_2	0.27	0.27	0.26	0.28	0.27	0.28	0.27
Al_2O_3	15.89	16.02	15.99	16.21	16.50	16.18	16.02
Fe_2O_3 (T)	1.55	1.58	1.57	1.77	1.77	1.64	1.61
MnO	0.041	0.033	0.038	0.043	0.062	0.039	0.049
MgO	0.81	0.82	0.77	0.77	0.72	0.82	0.66
CaO	1.83	2.17	2.16	1.97	2.06	2.18	2.17
Na_2O	5.62	5.67	5.56	5.87	6.08	5.75	5.55
K_2O	2.71	2.70	2.32	2.46	2.51	2.48	2.75
P_2O_5	0.10	0.12	0.11	0.11	0.11	0.12	0.12
Total	99.51	99.40	99.17	99.42	99.54	99.62	99.16
K_2O/Na_2O	0.48	0.47	0.41	0.42	0.41	0.43	0.49
ACNK	1.03	0.99	1.03	1.02	1.00	1.01	0.99
Trace elements							
Sc	1.4	1.5	—	2	2	—	2
V	15	15	—	15	14	—	13
Cr	11	21	—	12	9	—	18
Ni	2	1	3	1	—	3	—
Cu	—	4	5	10	14	7	10
Zn	43	66	55	49	44	59	51
Ga	19	18	19	20	19	19	18
Rb	107	86	106	117	195	136	153
Sr	647	910	836	722	759	881	653
Y	10	13	—	10	10	—	11
Zr	171	160	167	199	190	162	180
Nb	14	10	11	13	11	10	10
Ba	657	643	—	581	562	—	709
U	9	6	6	6	6	5	4
Th	25	17	17	21	21	13	4
La	52	—	—	59	—	—	—
Ce	81			87			
Nd	29			33			
Sm	5.2			6			
Eu	1.1			1.2			
Gd	3			3.4			
Dy	1.9			2			
Er	0.9			0.9			
Yb	0.8			0.8			
Lu	0.11			0.12			

the undeformed hornblende granodiorite phase exposed on the hillocks west of Mahoba railway station. Following these occurrences of older gneissic relicts in early Proterozioc intrusive granite phases, large exposures of older gneissic rocks (representing the Archaean crust) were delineated along Mahoba–Charkhari and Mahoba–Kabrai roads. One such large gneissic block (1 km long and 500 m wide) with associated metabasics is exposed around the Loda Pahar hillock, east of Kabrai, near the junction of Kabrai–Kanpur and Mahoba–Banda road, and is surrounded by intrusive biotite-bearing porphyritic granite phase. This hillock is 50 m high, the southern face of which has been extensively quarried for road construction material. On the northern slope of the Lodha Pahar, 1 m wide sulphide mineralized shear zone has been observed in the gneissic rocks which runs parallel to its northern contact with the biotite granite. Highly deformed and altered peridotites, and mafic schistose rocks are the host into which the trondhjemitic gneiss intrudes. Clear contact relationships among the various rock units in this area are masked by subsequent granitic activity and soil cover.

The Loda Pahar trondhjemitic gneiss (LPTG) is grey to pink, medium- to fine-grained gneiss with occasional phenocrysts of feldspar having a uniform mineral composition. The rock consists of quartz and plagioclase as dominant mineral with rare presence of K feldspar. Biotite altered to green chlorite is ferromagnesian mineral phase and is aligned parallel to the foliation of the rock. Accessory minerals include small euhedral crystals of Fe–Ti oxides, epidotes, apatite and zircon. All the samples exhibit typical granoblastic fabric modified by cataclastic deformation as evidenced by mortar texture, deformed and rapidly tapering lamellae in feldspars with locally developed myrmekites. The quartz, occurring as linear bands, shows undulose wavy extinction, recrystallized sutured grain contacts, with occasional granulated and polycrystallized fine-grained streaks parallel to the gneissosity.

Seven fresh and representative samples of Loda Pahar trondhjemitic gneiss were selected for geochemistry. All the major and trace elements were determined by XRF (Siemens SRS-3000 sequential X-ray spectrometer) and the rare-earth elements (REE) by AES-ICP at WIHG. The analytical methods for determining REE are similar to those proposed by Walsh[8]. The geochemical data of the LPTG are given in Table 1.

The chemical composition of the LPTG is quite uniform. The rock contains a restricted range of SiO_2 (69.45–70.69%), Al_2O_3 (15.89–16.5%), MgO (0.66–0.82%), .CaO (1.83–2.18%), Na_2O (5.55–6.08%) and K_2O (2.32–2.75%). The rock is poor in ferromagnesian elements ($Fe_2O_3 + MgO + TiO_2 < 5\%$) and has high Na_2O values relative to CaO and K_2O with consistent low K_2O/Na_2O (< 0.5) ratio. The sodic character of the LPTG is indicated by the K–Ca–Na triangular diagram (Figure 1), thus clearly reflecting trondhjemitic trend confirming to the definition of trondhjemite[9]. Further, the LPTG is classified as high-Al trondhjemites in view of its high Al_2O_3 content (> 15%) at 70% SiO_2 (refs 10–12). Its high-Al nature is also substantiated by lower ratio of K/Rb (106–260, av. 176) and Rb/Sr (av. 0.17), low Y (10–13 ppm), Cr (9–21 ppm, av. 14 ppm) and Nb (10–14 ppm), and high Sr (647–910 ppm) content. Similar values have been reported for Archaean high-Al TTG rocks from various continents[11]. Being uniform in major and trace element composition, two samples of LPTG were selected for REE analysis. The consistency of REE data and the highly fractionated $(La/Yb)_N$ (43, 49.6) with depleted Yb_N content (3.83) and concave upward shape at heavy REE (HREE) end without any significant Eu anomaly in the LPTG is remarkably similar to the typical high-Al TTG suite reported from the Archaean[11,13].

As mentioned earlier, the Archaean TTG at 70% SiO_2 level have been divided into geochemically distinct low-Al and high-Al types. This subdivision emphasizes the significant role of hornblende and its retention as a refractory phase in the source rock during partial melting, thereby increasing the Al_2O_3 content in the melt[14]. Various models for the origins of TTG parental magmas have been proposed by the earlier workers; however, the recent geochemical and experimental studies on the Archaean rocks convincingly support the idea of partial melting of garnet-bearing amphibolites[13].

The strongly fractionated REE pattern and HREE depletion in the LPTG magma (Figure 2) suggest that

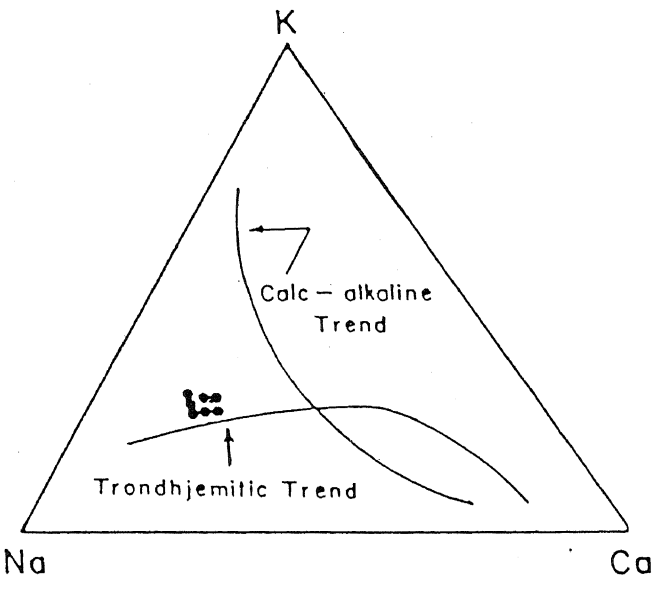


Figure 1. Ternary K–Na–Ca plot for the LPTG samples lying on the trondhjemite trend (from ref. 10).

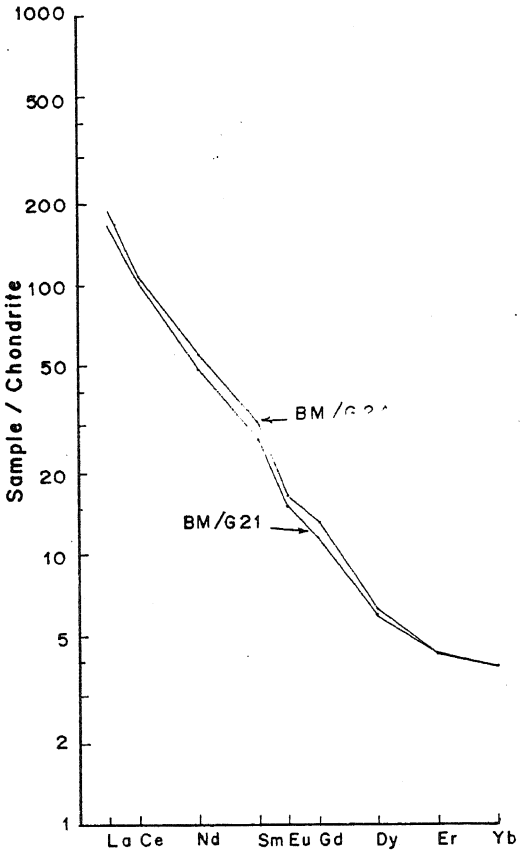


Figure 2. Chondrite normalized REE pattern of LPTG.

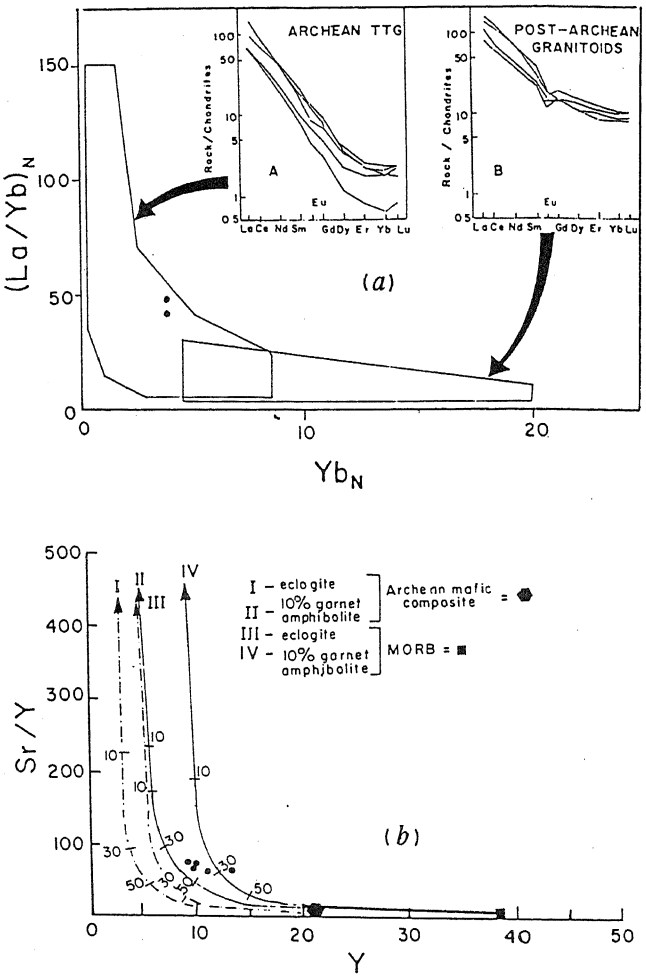


Figure 3. *a*, $(La/Yb)_N$ versus (Yb_N) diagram showing changes in REE contents with time (from ref. 13), the LPTG data plot in the Archean TTG field, *b*, Sr/Y versus Y diagram showing that the LPTG samples have an affinity with 10% garnet amphibolite source (from ref. 11).

hornblende and/or garnet may be the residual phase in the source. Further, the presence of garnet in the hornblende-rich source is implied by the metaluminous composition and the low normative corundum average values of 0.36%. The plots of the LPTG samples on $(La/Yb)_N$ vs Yb_N and Sr/Y vs Y diagram (Figure 3) fall in the Archaean TTG field and suggest the generation of magma by partial melting of the source rock comprised of 10% garnet amphibolite. Such a derivation of trondhjemitic magma is now well constrained by experimental and geochemical data by Martin[12,13], who suggested a three-stage process for their genesis: (i) initial production of a large volume of tholeiitic magma by partial melting of mantle; (ii) metamorphic transformation of this tholeiite into garnet amphibolite and their subsequent partial melting to give rise to the parental magmas of TTG, leaving hornblende + garnet + clinopyroxene + minor plagioclase; and (iii) fractional crystallization mainly of hornblende + plagioclase to produce the differentiated TTG suite. The highly fractionated REE pattern of the LPTG compares closely with the high-Al older metamorphic tonalitic gneiss (OMTG), Champua area of Singhbhum–Orissa craton (3.6 Ga)[15,16]; Markampara trondhjemite from Bastar craton (3.5 Ga)[17]; trondhjemitic gneiss from Satnur–Halagur–Sivasamudram area, Dharwar craton (2.9 Ga)[18] of the Indian Shield; trondhjemitic gneiss from Mount Edger batholith, Pilbara, Australian Shield (3.3 Ga)[19]; Ancient Gneissic Complex, Swaziland (3.0 Ga)[20,21] and Kivijarvi trondhjemitic gneiss from eastern Finland (2.8 Ga)[12]. In view of the remarkable geochemical similarity between the LPTG and other trondhjemitic rocks from different Archaean terrains, it is evident that the likely process of their formation may be the same.

To summarize, the presence of the typical high-Al trondhjemitic gneiss in association with other crustal components as relicts from the Kabrai area provides the first convincing evidence to qualify the Bundelkhand craton as a typical Archaean terrain similar to Singhbhum–Orissa, Bastar and Dharwar cratons of the Indian Shield.

1. Goodwin, A. M., *Precambrian Geology*, Academic Press, London, 1991, p. 666.

2. Radhakrishna, B. P., *J. Geol. Soc. India*, 1989, **34**, 1.
3. Rahman, A. and Zainuddin, S. M., *J. Geol.*, 1993, **101**, 413.
4. Basu, A. K., *Rec. Geol. Surv. India*, 1986, **117**, 61.
5. Crawford, A. R., *Can. J. Earth Sci.*, 1970, **7**, 91.
6. Sarkar, A., Trivedi, J. R., Gopalan, K., Singh, P. N., Singh, B. K., Das, A. K. and Paul, D. K., *Indian J. Earth Sci. CEISM Sem.*, 1984, 64.
7. Paul, D. K., 1994, personal communication.
8. Walsh, N., Buckley, F. and Barker, J., *Chem. Geol.*, 1981, **33**, 141.
9. Barker, F., in *Trondhjemite, Dacite and Related Rocks* (ed. Barker, F.), Elsevier, Amsterdam, 1979, pp. 1–12.
10. Barker, F. and Arth, J. G., *Geology*, 1976, **4**, 596.
11. Drummond, M. S. and Defant, M. J., *J. Geophy. Res.*, 1991, **B95**, 21503.
12. Martin, H., *J. Petrol.*, 1987, **28**, 921.
13. Martin, H., *Lithos*, 1993, **30**, 373.
14. Zen, E., *J. Petrol.*, 1986, **27**, 1095.
15. Saha, A. K., Ray, S. L., Ghosh, S., Mukhopadhyaya, K. and Dasgupta, D., *Mon. Indian Soc. Earth Sci.*, 1984, 1.
16. Saha, A. K., *Mem. Geol. Soc. India*, 1994, **27**, 341.
17. Sarkar, G., Corfu, F., Paul, D. K., McNaughton, N. J., Gupta, S. N. and Bishnui, P. K., *Precambrian Res.*, 1993, **62**, 127.
18. Mahableshwar, B., Jayananda, M., Peucat, J. J. and Swamy, N. S., *J. Geol. Soc. India*, 1995, **45**, 33.
19. Jahn, B. M., Glikson, A. Y., Peucat, J. J. and Hickman, A. H., *Geochim. Cosmochim. Acta*, 1981, **45**, 1633.
20. Hunter, D. R., Barker, F. and Millard, H. T., Jr., *Precambrian Res.*, 1987, **7**, 105.
21. Hunter, D. R., Barker, F. and Millard, H. T., Jr., *Precambrian Res.*, 1984, **24**, 131.

ACKNOWLEDGEMENTS. We are grateful to Dr P. P. Khanna and Mr N. K. Saini for analytical work. The financial assistance provided by the Department of Science and Technology, Government of India, under Deep Continental Studies Programme is thankfully acknowledged.

Received 16 May 1995; revised accepted 27 May 1995

Hypothalamic γ-glutamyl transpeptidase activity in rats following morphine and naloxone treatment

P. Sreenivasan and E. Vijayan

School of Life Sciences, Pondicherry University, Pondicherry 605 014, India

Hypothalamic γ-glutamyl (γ-GT) transpeptidase activity was measured in pubertal and sexually mature male rats. Rats were given subcutaneous injection of morphine and naloxone and the enzyme activity was measured at different time intervals. Morphine had an inhibitory effect on the enzyme activity in both the age groups. Naloxone, an opioid receptor antagonist, had a stimulatory effect on γ-GT activity in pubertal rats while it had no effect in sexually mature rats. This suggests that opiates exert an inhibitory effect on glutathione utilization through γ-GT cycle. Physiological significance of this finding with reference to pituitary function appears to be the suppression of gonadotropin secretion.

GAMMA-glutamyl transpeptidase (γ-GT) is a widely distributed enzyme that functions in the utilization of glutathione by transferring the γ-GT moiety of the tripeptide to amino acid or peptide acceptors. This enzyme together with others plays a major role in the synthesis and degradation of glutathione via the γ-GT cycle[1]. Gamma-GT is a membrane-bound enzyme present in abundance in epithelia of many organs[2,3]. The substrates for the enzyme are glutathione (GSH), oxidized glutathione (GSSG), S-substituted glutathione and other γ-GT compounds. Glutathione's function is to maintain protein thiol groups which may be required for catalysis and is involved in protein assembly and degradation. Glutathione also helps in protecting protein and cell membranes against peroxides and free radicals. Glutathione has also been reported to play a modulatory role on anterior pituitary hormone release[4]. It is probable that glutathione interacts with neuropeptides such as GnRH, opioid peptides and other regulatory elements at the hypothalamic level to elicit an endocrine response. The present study was designed to evaluate whether hypothalamic γ-GT transpeptidase activity responds to treatment with morphine sulphate – an opiate agonist – and naloxone – an opioid receptor antagonist – and thereby affects glutathione metabolism in pubertal and sexually mature rats.

All the experiments were done in colony-bred rats, derived from Wistar strain, purchased from the Animal Facility of Jawaharlal Institute of Postgraduate Medical Education and Research (JIPMER), Pondicherry, and Kind Institute of Preventive Medicine, Madras. Forty (pubertal) and 60 (sexually mature) days old male rats were used in this study. They were maintained under controlled conditions of light (14 h light and 10 h dark) and temperature with free access to drinking water and standard rat pellets.

Morphine sulphate, γ-GT paranitroanilide, glycyl–glycine and naloxone were obtained from Sigma Chemical Co., St Louis, USA. Naloxone (2 mg/kg body weight, dissolved in 0.9% saline), morphine sulphate (2 mg/kg body weight dissolved in 0.9% saline) were injected subcutaneously. All the doses employed are known to produce physiological responses[5,6]. Rats were sacrificed at 1 and 4 h after injection and hypothalamus was quickly dissected out and processed for estimation of γ-GT transpeptidase. The enzyme activity was estimated according to the method of Tate and Meister[7] as standardized by Pasha and Sadasivudu[8]. This method has been widely used to measure the enzyme activity in

choroid plexus and other areas of brain[8]. The assay mixture consisted of 80 µmol of Tris-HCl buffer (pH 9.0), 150 µmol of NaCl, 40 µmol of glycyl glycine, 5 µmol of γ-GT paranitroanilide and 0.4 ml of 2% homogenate (in 0.25 M sucrose). After 30 min incubation at 37°C the reaction was terminated by the addition of 2 ml of 10% acetic acid. After centrifugation at 5000 rpm for 20 min the absorbance of the clear supernatant was read at 410 nm. The enzyme activity was expressed as µmol of paranitroaniline liberated per gram weight of tissue per hour with reference to a standard graph plotted using different concentrations of paranitroaniline and treating it with acetic acid.

Statistical analysis was done using analysis of variance, followed by Duncan's new multiple range test.

γ-GT activity ranged from 44.14 ± 0.64 in 40-day-old to 52.96 ± 1.91 µmol *p*-nitroaniline released/g wt tissue/h in 60-day-old rats. Administration of morphine to 40-day-old rats did not modify the enzyme activity at 1 h, whereas the activity was reduced at 4 h ($P < 0.05$) after injection. In contrast, naloxone treatment produced an increase in the γ-GT activity at 1 h ($P < 0.05$) while there was no change at 4 h (Table 1). Morphine treatment to 60-day-old rats reduced hypothalamic γ-GT activity significantly both at 1 h ($P < 0.01$) and at 4 h ($P < 0.05$) after injection. Naloxone, however, caused no change in the enzyme activity either at 1 or 4 h in 60-day-old rats (Table 2).

The present results show that morphine, a natural opiate agonist and naloxone, an opiate receptor antagonist, produce changes in hypothalamic γ-GT transpeptidase activity. Though morphine sulphate failed to modify γ-GT at 1 h, it produced significant reduction at 4 h in 40-day-old rats. Administration of naloxone, on the other hand, produced an opposite effect. Earlier studies showed that glutathione content in hypothalamus reached a peak at puberty. Further, intraventricular administration of glutathione resulted in a significant increase in plasma FSH and a reduction in LH levels in ovariectomized oestrogen–progesterone primed rats[3]. Such an effect of glutathione on gonadotropin levels together with the observation of pubertal peak in glutathione content suggested an interaction of glutathione with the hypothalamic–pituitary axis. The increase in γ-GT activity in pubertal rats is consistent with the above observation. Further, it is significant in consonance with the recently held view that attenuation of opioid inhibition at the time of puberty in male and female rats is one mechanism for the onset of puberty[9,10]. Increased utilization of glutathione through γ-GT cycle maybe one mechanism for increased secretion of gonadotropins[3]. In sexually mature rats naloxone did not modify γ-GT activity, suggesting noninvolvement of glutathione utilization in endogenous opioid-mediated hormone response in these rats. Morphine sulphate had a marked inhibitory effect on hypothalamic γ-GT activity in both 40- and 60-day-old rats. Administration of morphine is reported to be inhibitory to gonadotropin secretion[5,6,11]. Available evidences suggest that glutathione utilization by the γ-GT cycle is involved in gonadotropin release[4,12]. A decrease in γ-GT activity following morphine treatment suggests that this may be one of the mechanisms of morphine-induced suppression of gonadotropin secretion. Morphine is known to regulate release of other pituitary hormones as well[5]. It may do so by altering the release of neurotransmitters at the hypothalamic level. The physiological significance of the effect of morphine on glutathione utilization through the γ-GT cycle on pituitary hormone release is not clear. The formation of γ-GT derivatives such as γ-GT dopamine may have functional significance as dopamine is known to influence release of a number of hormones from the pituitary[13–15].

Since opiates of other classes (prodynorphine and proenkephalin) are less potent in eliciting a response on gonadotropin release as compared to morphine, they were not used in the present study. However, the present results suggest that hypothalamic γ-GT transpeptidase activity is altered under the influence of morphine and reversed by opioid receptor antagonist naloxone. The exact mechanism and physiological significance of this response is yet to be understood.

The data presented suggest that opioid peptides, have an inhibitory effect on glutathione utilization through the γ-GT cycle. In pubertal rats, opioid receptor antagonist naloxone stimulated the γ-GT activity, indicating endogenous opioid-mediated suppression of glutathione utilization. This is significant with respect to onset of

Table 1. Hypothalamic γ-glutamyl transpeptidase activity in 40-day-old rats

Treatment	1 h	4 h
Saline control	44.14 ± 0.64	43.21 ± 0.8
Morphine sulphate	42.98 ± 1.2	34.62 ± 1.2*
Naloxone	52.10 ± 4.3*	46.13 ± 2.4

*$P < 0.05$.
Values are in µmol *p*-nitroaniline released/g wt tissue/h.
Values represent mean ± SE of 6 rats in each group.

Table 2. Hypothalamic γ-glutamyl transpeptidase activity in 60-day-old rats

Treatment	1 h	4 h
Saline control	52.96 ± 1.91	51.40 ± 1.0
Morphine sulphate	44.47 ± 0.5**	46.10 ± 2.5*
Naloxone	49.93 ± 2.2	55.80 ± 1.3

*$P < 0.05$.
**$P < 0.01$; significantly different from control value.
Values are in µmol *p*-nitroaniline released/g wt tissue/h.
Values represent mean ± SE of 6 rats in each group.

puberty as glutathione level in hypothalamus is reported to reach a peak at puberty. Suppression of γ-GT transpeptidase activity by morphine in both the age groups suggests that this may be one of the mechanism(s) involved in suppression of gonadotropin secretion.

1. Meister, A. and Tate, S. S., *Annu. Rev. Biochem.*, 1976, **45**, 559–604.
2. Tate, S. S., Ross, L. L. and Meister, A., *Proc. Natl. Acad. Sci. USA*, 1973, **70**, 1447–1449.
3. Pasha, V. K. and Vijayan, E., *Brain Res. Bull.*, 1989, **22**, 617–619.
4. Pasha, V. K. and Vijayan E., *Biochem. Int.*, 1990, **21**, 209–217.
5. Meites, J., Bruni, J. F., Van Vugt, D. A. and Smith, A. F., *Life Sci.*, 1979, **24**, 1325–1336.
6. Sreenivasan, P., Ph D thesis, Pondicherry University, Pondicherry, 1994.
7. Tate, S. S. and Meister, A., *J. Biol. Chem.*, 1974, **249**, 7593–7602.
8. Pasha, V. K. and Sadasivudu, B., *Neurosci. Lett.*, 1984, **46**, 209–212.
9. Bhanot, R. and Wilkinson, M., *Endocrinology*, 1983, **113**, 596–603.
10. Wilkinson, M. and Kathleen, M. L., in *Brain Opioid Systems in Reproduction* (eds Dyer, R. G. and Bicknell, R. J.), Oxford University Press, New York, 1989, pp. 70–91.
11. Kalra, S. P., Allen, L. G. and Kalra, P. S., in *Brain Opioid Systems in Reproduction* (eds Dyer, R. G. and Bicknell, R. J.), Oxford University Press, New York, 1989, pp. 95–111.
12. Pasha, V. K. and Vijayan, E., *Biochem. Int.*, 1992, **26**, 7–15.
13. Tsuji, M., Matsuoka, Y. and Nadajima, T., *J. Neurochem.*, 1974, **29**, 633–637.
14. Ichinose, H., Togari, A., Suzuki, A. and Nagatsu, T., *J. Neurochem.*, 1987, **49**, 928–932.
15. Vijayan, E. and McCann, S. M., *Neuroendocrinology*, 1978, **25**, 221–235.

ACKNOWLEDGEMENTS. We thank University Grants Commission for a research grant to EV. PS was a recipient of JRF and SRF from UGC.

Received 9 February 1995; revised accepted 11 September 1995

Inhibition of biogas production by isobutyric and isovaleric acids and their precursors

M. K. Sharma* and M. M. Simlot†

Department of Biochemistry, Rajasthan College of Agriculture, Udaipur 313 001, India
*Present address: Agricultural Research Station, Mandore, Jodhpur 342 004, India
†Present address: 24 Subhashnagar, Opp. B. N. College, Udaipur 313 001, India

Isobutyric and isovaleric acids inhibit the production of biogas in the anaerobic fermentation of cowdung slurry at a concentration of 1000 ppm. Amino acids leucine and valine, whose metabolic products are isovaleric and isobutyric acids, respectively, when added to the fermenting media did inhibit the biogas production, but not to the extent free acids do, as some of the amino acids could have been utilized for protein synthesis.

METHANE gas production under anaerobic fermentation of agricultural wastes and animal wastes is a complex multistep system involving different microorganisms, which first degrade cellulose and turn the product into methane with the help of methanobacterium present in the medium[1–3]. Many investigators have reported about the importance of volatile acids in the production of methane gas from cattle wastes. Varel *et al.*[4] observed that, under conditions of optimum methane production from cattle wastes, total volatile acid is the first to accumulate in large amounts by increasing the total volatile solids in the feed. The additive effect of addition of water hyacinth[5–7] and poultry waste[8] to cowdung enhanced the volatile fatty acid content. This was further confirmed using a two-phase fermenting system[9]. Hobson and Shaw[10] reported that volatile fatty acids acetate and butyrate were not inhibitory to methane production by *Methanobacterium formicium*, but propionate was under certain conditions. Kalle and Menon[11] reported that methanogenesis was inhibited by branched-chain fatty acids like isobutyric and isovaleric acids below ambient temperature. However, Kalle *et al.*[12] reported that normal mixed cellulolytic culture produces straight-chain fatty acids, which could be due to the presence of isovaleric-degrading bacteria present in marine sediment and sewage

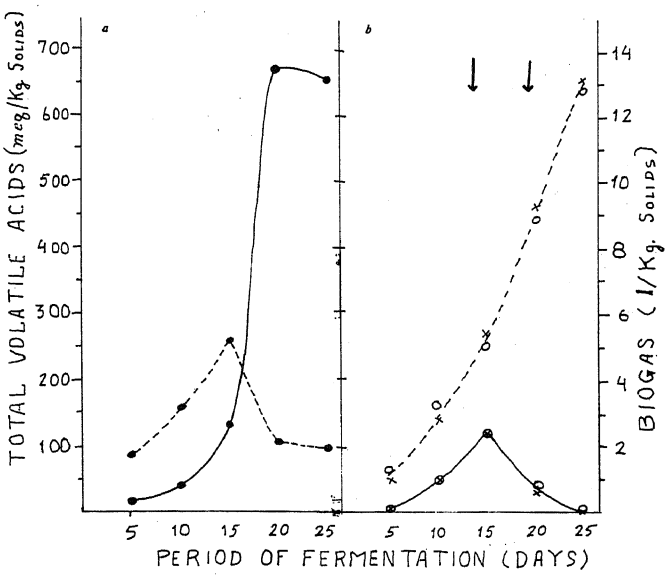


Figure 1. Effect of addition of isobutyric or isovaleric acid to the fermenting medium. Isobutyric acid (IBA) or isovaleric acid (IVA) was added on the 15th and 20th days of fermenting to a final concentration of 1000 ppm and the total amounts of volatile acids and biogas produced were measured. *a*, Control: ●----● total volatile acids; ●——● biogas produced. *b*, IBA or IVA added: O----O IVA, total volatile acids; O——O IVA, biogas produced; ×----× IBA, total volatile acids; ×——× IBA, biogas produced.

Table 1. Biogas production from cowdung in the presence of valine or leucine

Treatment/ Days	5	10	15	20	25
	(l/kg solid)				
Control	0.22	1.36	3.49	12.33	12.11
+ valine (2000 ppm)	0.15	1.05	2.89	11.06	10.94
+ leucine (2000 ppm)	0.17	1.24	3.27	11.42	11.15
SEM ±	0.013	0.050	0.074	0.061	0.051
CD (5%)	0.046	0.176	0.259	0.212	0.178
CD (1%)	0.070	0.267	0.392	0.321	0.270

NB: Values are means of four replicates.

Table 2. Total volatile fatty acids in the medium in the presence of valine and leucine

Treatment/ Days	5	10	15	20	25
	(meq/100 g solid)				
Control	7.79	15.53	25.86	14.20	13.19
+ valine (2000 ppm)	7.86	16.40	29.20	18.40	17.33
+ leucine (2000 ppm)	7.06	15.66	27.46	15.86	15.33
SEM ±	0.254	0.139	0.263	0.213	0.313
CD (5%)	0.880	0.482	0.912	0.739	1.084
CD (1%)	1.334	0.731	1.755	1.120	1.642

NB: Values are means of four replicates.

sludge[13]. Since the presence of inhibitors may also affect the rate of biogas production, the present study is directed towards the effect of branched-chain fatty acids and their precursors on biogas production.

Fresh cowdung was collected from the cattle farm of Rajasthan College of Agriculture, Udaipur. Volatile fatty acid in the fermenting media was estimated according to the method of Scarisbrick[14], by acidifying the sample with phosphoric acid to pH 2.5, diluting with equal volume of water and centrifuging to remove the solid material. Biogas production was measured by collecting the gas in an apparatus designed in this lab (Simlot, unpublished), which was connected to a digestor. The digestor contained 7.5% slurry of cowdung (total volume 400 ml) and was allowed to digest at 35 ± 2°C. The gas was collected for 5 days and measurements made every 5th day.

The production of biogas depends on the medium, temperature and other anaerobic conditions of the medium and normally begins between days 3 and 12 at 55°C in cowdung fermentation[4]. As shown in Figure 1, in this case it started after 4 days and reached the maximum production of 13.46 l/kg solid on the 20th day in the control. Thereafter, the biogas production started declining to the level of 13.22 l/kg solid on the 25th day. The amount of total volatile fatty acids rose, reaching an optimum level of 262.9 meq/kg solid on the 15th day of fermentation and then dropping suddenly to 105.8 meq/kg solid on the 20th day. These observations about sudden spurt in biogas production from 15th to 20th day and concomitant drop in the total volatile fatty acid content confirms the earlier report that biogas production is dependent on the formation of volatile fatty acids[3]. On addition of branched-chain fatty acids like isobutyric and isovaleric acids (1000 ppm each) to the fermenting medium on the 15th day, when the biogas production had started picking up, it was found that gas production had gone down considerably, while at the same time the total volatile fatty acid content, instead of coming down, had started going up. Another dose of these acids on the 20th day had totally blocked the production of biogas. These results confirm the observations of Kalle and Menon[11] regarding the inhibitory effect of these acids on methane production.

Leucine and valine through a number of intermediate steps are metabolized to isovaleric and isobutyric acids, respectively. The presence of these amino acids in the fermenting medium can be used in two ways – either for the synthesis of new proteins or enzymes, or the excess may metabolize to isobutyric and isovaleric acids. The conversion of leucine and valine to isovaleric and isobutyric acids will inhibit the methanogenesis. Table 1 gives the volume of the biogas produced when leucine or valine was added to the fermenting medium. Initially, there was not much change in the amount of biogas produced in comparison to the control. But after the 15th day, biogas production was significantly reduced as compared to the control and remained at the lower level thereafter. The total volatile fatty acid content (Table 2), however, increased significantly over the control. Thus, these results indicate that the two amino acids are partially metabolized to their corresponding fatty acids, i.e. isovaleric and isobutyric acids, respectively, with the result that biogas production is inhibited.

1. Chen, Y., Varel, V. H. and Hashimoto, A. G., *Ind. Eng. Chem. Prod. Res. Develop.*, 1986, **10**, 471.
2. Zeikus, J. G., *Annu. Rev. Microbiol.*, 1980, **34**, 423.
3. Mackie, R. I. and Bryant, M. P., *Appl. Environ. Microbiol.*, 1981, **41**, 1363–1373.
4. Varel, V. H., Issacson, H. R. and Bryant, M. P., *Appl. Environ. Microbiol.*, 1977, **33**, 298–307.
5. Deshpande, P., Sarnaik, S., Godbole, S. H. and Wagle, P. M., *Curr. Sci.*, 1979, **48**, 490–492.
6. Simlot, M. M., National Symposium on Recycling Residue of Agriculture Industry, 28–29 March 1980.
7. Kumar, S. and Jain, M. C., *Res. Ind.*, 1985, **30**, 91–94.
8. Singh, R., Jain, M. K. and Tauro, P., *HUA J. Res.*, 1982, **12**, 107–123.
9. Lo, K. V. and Liao, P. H., *Energy Agric.*, 1986, **5**, 81–88.
10. Hobson, P. N. and Shaw, B. G., *Water Res.*, 1976, **10**, 849–852.
11. Kalle, G. P. and Menon, K. K. G., *J. Biosci.*, 1984, **6**, 315–324.

12. Kalle, G. P., Nayak, K. K. and Desa, C., *J. Biosci. (Bangalore)*, 1985, **9**, 137–144.
13. Stieb, M. and Schink, B., *Arch. Microbiol.*, **144**, 291–295.
14. Scarisbrick, R., in *Modern Methods of Plant Analysis* (eds Paech, K. and Tracey, M. V.) Springer, Berlin, 1955, vol. II, pp. 455–458.

Received 9 March 1994; revised accepted 14 July 1995

Production of wheat haploids through embryo rescue from wheat × maize crosses

N. S. Bains, Jaswinder Singh, Ravi* and S. S. Gosal**

Department of Plant Breeding, *Department of Genetics and **Biotechnology Centre, Punjab Agricultural University, Ludhiana 141 004, India

Appropriate hormone treatment using 2,4-D and gibberellic acid led to a high frequency (34.17%) of embryo formation in crosses of field-grown wheat strains with maize. Twenty-five days after culture on MS medium supplemented with casein (2000 mg/l) and amino acids, 30.95% of the embryos showed plant formation, 44.05% resulted in callus-like structures and the remaining failed to respond. The haploid nature of some of the regenerated plants was confirmed cytologically. Implications for wheat breeding are discussed.

WHEAT is a premier food crop of worldwide importance. It is also a crop where conventional plant breeding has paid rich dividends, as epitomized by the green revolution. However, the current impasse in yield levels calls for application of newer techniques. A major hindrance to rapid genetic turnover in self-pollinated crops, such as wheat, is the inordinately long time (generally 8–10 years) it takes to develop stable, homozygous and ready-to-use material from a fresh cross. In this context, anther/microspore culture has often been proposed as a method for producing instant homozygous lines via haploidy. In wheat, wide hybridization followed by chromosome elimination serves as an alternative route to haploidy, e.g. wheat × *Hordeum bulbosum* and wheat × *Zea mays* crosses. The wheat × *Z. mays* cross is a relatively new system which offers distinct advantages such as freedom from genotypic specificity, the bane of anther culture approach as well as the wheat × *H. bulbosum* system (which is restricted to wheat genotypes carrying the kr crossability genes). Moreover, the system is less prone to gametoclonal variation owing to the absence of a dedifferentiated phase.

Zenkteler and Nitzsche[1] were the first to report microscopic, early-stage embryos in crosses between wheat and maize. However, the frequency of embryo formation was not specified, nor was evidence presented for their hybrid origin. Intrigued by this report, Laurie and Bennett[2] set out to study early post-pollination events in wheat and maize crosses under a CIMMYT project. They demonstrated the presence of both wheat and maize chromosomes in the zygotes and observed further that maize chromosomes were eliminated during initial cell divisions. Endosperm development ceases early or never occurs and embryos fail to develop to a size that can be readily rescued. Two years later, they reported[3] the recovery of the first haploid plants using the wheat × maize system by employing *in vitro* culture of wheat spikelets, 2 days after pollination. A method that bypassed spikelet culture was devised by Suenaga and Nakajima[4], who injected 2,4-D solution (100 ppm) into the uppermost internodes of wheat stems to sustain embryo growth on the plant itself, till the appropriate stage for embryo rescue.

The present study was conducted as a preliminary assessment of the applicability of the system to indigenous material grown under field conditions and to judge its potential as a plant breeding tool. The wheat material used represented advanced, agronomically relevant homozygous lines. (For actual plant breeding purposes, however, heterozygous material would be required.) Wheat ears were emasculated 3–5 days prior to anthesis by opening up the glumes and removing the anthers, unlike the conventional practice, which involves the cutting of the glumes to expose the androecium. Freshly collected pollen from a winter maize variety, Partap 1, was used for pollinations, 3–5 days after emasculation. The pollinated tillers were administered injections of 2,4-D solution (125 ppm) into the uppermost internode, daily in the evening, for three days after pollination. On the fourth day the ears were momentarily dipped in a solution of 2,4-D (100 ppm) + gibberellic acid (75 ppm). On an average, seed-like structures (Figure 1 *a*) were seen to develop in 86.71% of the wheat ovaries pollinated with maize (Table 1). Closer examination of the ears revealed that ovaries which failed to develop into seed-like structures were damaged by severe fungal infection. The presence of a large amount of maize pollen in the floret and moisture provided by hormonal (2,4-D) ear dip might have favoured this fungal growth. Thus, seed-like structures develop in all the viable florets. Obviously, all such seeds are not the result of fertilization and are probably formed due to hormone induced enlargement. Thus, actual seed setting could be assessed only after dissecting the seed-like structures and observing the presence/absence of embryos. Twenty developing seeds each of six genotypes 15 days after pollination with maize were specifically dissected for this purpose. No visual differentiation seemed to exist between seeds that carry an embryo

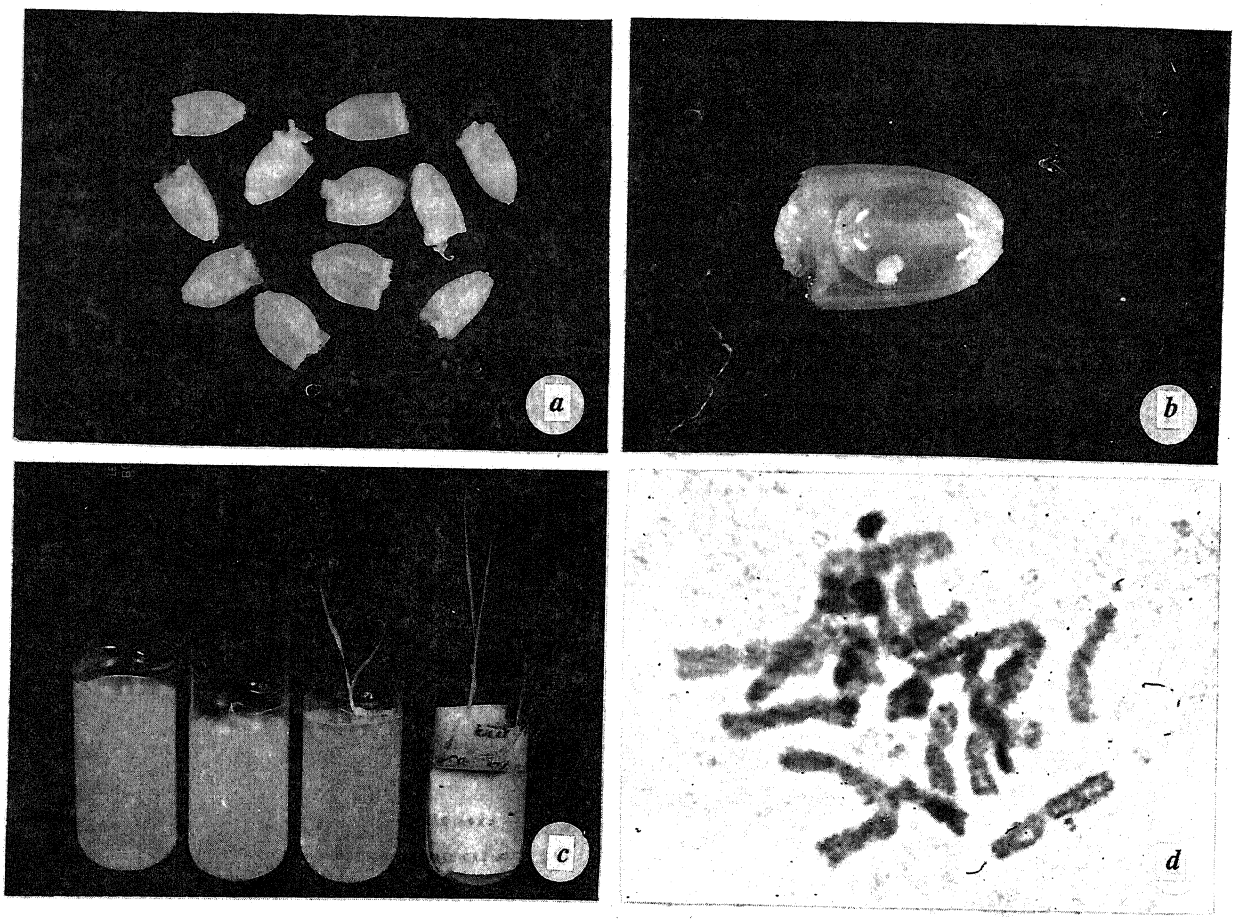


Figure 1. Wheat × maize crosses: *a*, seed-like structures observed in the crossed florets; *b*, haploid embryo floating in the watery endosperm; *c*, regeneration of haploid plants; *d*, root tip cell from embryo-derived plant showing haploid chromosome number ($2n = 21$).

Table 1. Seed setting, embryo formation and response to embryo culture in wheat × maize crosses

Wheat genotype used as female parent	Percentage of pollinated florets showing formation of seed-like structures (15 days after pollination)	Percentage of seed-like structures showing embryo formation* (15 days after pollination)	Percentage of cultured embryo (25 days after culture) showing		
			plant regeneration	callus formation	no response
PBW 299	91.86 (79/86)**	20.00 (4/20)	12.50 (1/8)	50.00 (4/8)	37.50 (3/8)
PBW 343	86.73 (85/98)	—	10.00 (1/10)	50.00 (5/10)	40.00 (4/10)
PBW 352	82.41 (89/108)	55.00 (11/20)	8.33 (1/12)	66.67 (8/12)	25.00 (3/12)
PBW 362	93.75 (60/64)	5.00 (1/20)	—	—	—
PBW 373	80.36 (90/112)	—	41.67 (5/12)	50.00 (6/12)	8.33 (1/12)
PBW 377	93.40 (99/106)	5.00 (7/20)	25.00 (2/8)	37.50 (3/8)	37.50 (3/8)
PBW 391	82.73 (91/110)	—	41.67 (5/12)	16.67 (2/12)	41.67 (5/12)
PBW 396	87.50 (98/112)	60.00 (12/20)	54.54 (6/11)	45.45 (5/11)	0.00 (0/11)
PBW 398	86.06 (105/122)	30.00 (6/20)	45.45 (5/11)	36.36 (4/11)	18.18 (2/11)
Mean	86.71 (796/918)	34.17 (41/120)	30.95 (26/84)	44.05 (37/84)	25.00 (21/84)

*Based on dissection of 20 seeds per genotype under a stereomicroscope.
**Values in parenthesis are the actual numbers of observations.

and those that do not. The frequency of embryos ranged from 5 to 60% (Table 1). Small sample size makes it difficult to say whether this denotes genotypic specificity. However, remarkably high overall frequency of 34.17% indicates the scope for applying this system to plant breeding situations.

The second important factor for the success of this system is the frequency of plant emergence from the cultured embryos. For culturing, 15-day-old (days after pollination with maize) seed were surface-sterilized in 70% alcohol for 30 s followed by a 10 min treatment with 0.15% mercuric chloride and three rinses in sterile water. Embryos were excised aseptically under a stereo microscope and cultured in a medium containing MS basal salts supplemented with casein (2000 mg/l), L-glutamine (400 mg/l), L-alanine (50 mg/l), L-cysteine (20 mg/l), L-arginine (10 mg/l), L-leucine (10 mg/l), inositol (100 mg/l) and sucrose (50 g/l). The embryos were atypical in shape, having microcalli-like appearance, and were found floating in the watery endosperm (Figure 1 *b*). The cultures were incubated in total darkness for 10 days at 25 ± 1°C and then shifted to 16 h/8h light/dark regime.

The response of the embryos was recorded 25 days after the culture and is shown in Table 1. Out of the total 84 embryos cultured, 26 showed plant emergence (Figure 1 *c*), giving a frequency of 30.95% averaged over genotypes. A considerable proportion (44.05%) of the embryos formed callus-like structures probably due to the carryover effect of 2,4-D treatments. Transfer to regeneration medium in the subsequent culture can result in plant formation from these embryo-derived calli. About 25% of the embryos showed no response. Further experimentation for fine-tuning of the medium composition can greatly enhance the recovery of haploid plants.

For confirming the haploid nature of the regenerated plants, mitotic chromosome analysis of root tips was conducted. Some randomly taken plants were analysed. The roots were pretreated in a saturated solution of α-bromonaphthalene for 2 h and fixed in Carnoy's solution for 24 h. The roots were then hydrolysed in 1 N HCl at 60°C for 10 min and stained with 2% aceto-carmine. The dividing cells carried a chromosome complement of 21 chromosomes (Figure 1 *d*) in all the plants as revealed by light microscopy. The normal somatic chromosome number of hexaploid wheat is 42.

The wheat × maize system holds promise for revolutionizing wheat breeding methodology. Homozygosity is approached in one step, thereby circumventing a period of 6–8 years in the breeding cycle. This system can be very valuable for situations requiring immediate genetic amelioration, e.g. incorporation of resistance to major crop diseases. In addition, doubled haploids from crosses of divergent wheat parents is a short cut in obtaining recombinant inbreds, which are very useful for construction of genetic linkage maps using molecular markers. Haploids can also be useful for induction and recovery of mutants as well as for basic studies involving genomic relationships.

1. Zenkteler, M. and Nitzsche, W., *Theor. Appl. Genet.*, 1984, **68**, 311–315.
2. Laurie, D. A. and Bennett, M. D., *Can. J. Genet. Cytol.*, 1986, **28**, 313–316.
3. Laurie, D. A. and Bennett, M. D., *Theor. Appl. Genet.*, 1988, **76**, 393–397.
4. Suenaga, K. and Nakajima, K., *Plant Cell Rep.*, 1989, **8**, 263–266.

Received 24 May 1995; revised accepted 7 August 1995

Molecular Genetics of Bacterial Pathogenesis: A Tribute to Stanley Falkow. Miller, V. L., Kaper, J. B., Portnoy, D. A. and Isberg, R. R. (eds). American Society for Microbiology, ASM Press, Washington, DC 20005, USA. 1994. XL + 529 pp. Price: £ 59.50.

This volume is meant to commemorate the 60th birthday of Stanley Falkow, who, with his enunciation of 'molecular koch's postulates', laid the foundations of molecular genetic analysis of bacterial pathogens. 'Legacy of Stanley Falkow' by Kaper and Miller brings out his extraordinary talents as a researcher matched by his exceptional skills as a teacher. In 'A look through retroscope' Falkow himself recounts the origins of his interests in bacterial pathogenicity, which started after reading *Microbe Hunters* at a very young age, and culminated in unravelling several of the molecular underpinnings of bacterial pathogenicity.

The main text of the book is divided into six parts.

Part I – 'Retrospective look at early advances' – is a historical account of how Falkow's work on extrachromosomal elements, especially R factor, increased the awareness of the scientific community that plasmids were not a curiosity of geneticists but rather had far-reaching practical implications, viz. role in bacterial pathogenesis and antibiotic resistance (Chapters 1, 2). There are reviews on the role of Tn3 in the spread of antibiotic resistance (Chapters 3, 4) and application of plasmid profile analysis to trace the epidemiology of pathogenic bacteria, which is now a burgeoning field of 'molecular epidemiology' (Chapter 5).

Colonization of the host surfaces by pathogenic bacteria represents the first and an almost indispensable step in the initiation of infectious disease process. Thus, an interest in colonization factors, both for studying the bacterial pathogenicity and for designing interventional strategies, cannot be overemphasized. Part 2 – 'Adhesins' – discusses molecular genetics of an array of colonization factors, viz. adherence factors of urinary tract pathogens (Chapter 6), *E. coli* type 1 pili (Chapter 7), pilus antigenic variation (Chapter 8) and type 4 pili (Chapter 9). These chapters not only depict the state-of-the-art of the subject but also relate the chequered path of the development of each of these areas.

Until recently, adhesins were seen as appendages passively mediating attachment of pathogens to single receptor on the host or target cells. It is, however, becoming clear that adhesins are much more complex structures in having multiple binding sites and their function is not limited to passive binding but extends to signal transduction in host/target cells. To carry out such functions adhesins mimic host molecules whose natural functions involve adherence and signal transduction. Also, co-operation exists between different adhesins in establishing an infection which is reflected in the organization and regulation of their genes. These aspects have been discussed in relation to adhesins of *Bordetella pertussis* (Chapter 10), *Haemophilus influenzae* (Chapter 11) and E-PEC (Chapter 12). Discussion on murine colonic hyperplasia, a naturally occurring disease of mice caused by *Citrobacter freundii* biotype 4280, has obvious relevance to the study of human proliferative bowel disorders (Chapter 13).

Intracellular bacteria, whether obligate or facultative, present enormous challenges to scientific and medical community. Part 3 – 'Molecular and cellular biology of intracellular bacteria' – present a detailed account of the molecular underpinnings of virulence determinants of intracellular bacteria and the cell biology of their interaction with host cells. Exhaustive genetic approaches have been used to identify *Salmonella* virulence genes. The various genes under investigation are: *spv* (*S*almonella *p*lasmid *v*irulence), *spa* (*s*urface *p*resentation of *a*ntigen), *inv* (*inv*asion), *pag* (*Pho-P a*ctivating *g*enes) (Chapter 14). Spectacular events in the host cell which accompany the invasion of *Yersinia* (Chapter 15) and *Salmonella* (Chapter 16) into the cell are also discussed.

Studies on the molecular aspects of the growth of *Legionella pneumophila* within the phagocytic cells have been presented to provide a comprehensive model by which pathogens may manipulate the host cells to suit to their intracellular lifestyle (Chapter 17). The host response to *Listeria monocytogenes*, an important food-borne pathogen, has been reviewed as a paradigm of cellular immune response to intracellular pathogens (Chapter 18). The emphasis of the review is on the cell biology of this infection, i.e. pathogen's entry into the host cell, its escape from the phagocytic vacuole using listeriolysin O (LLO) and the role of actin-driven movement in intercellular spread. The molecular genetics of chlamydial pathogenesis are currently under intensive investigation. In this context, the recently proposed chlamydial adhesins, viz. major outer membrane protein (MOMP), chlamydial cytoadhesin (CCA), Hsp-70 and another outer membrane protein (omp-2) have been reviewed. The potentialities of some of these antigens for the development of a vaccine have also been discussed (Chapter 19).

Part 4 – 'Extracellular structures and products' (Chapters 20–26) – is a critical account of the molecular genetics of the most obvious virulence determinants, viz. capsules (*E. coli* K-1 and Group B streptococci – type III), toxins (pertussis toxin and *E. coli* verotoxins), *E. coli* hemolysin, flagella (*Campylobacter* spp.) and yops (*Yersinia* spp.). Detailed information on the molecular genetics of these products has been gained primarily because these virulence factors have been most amenable to tools of molecular biology. The biosynthesis and assembly of these virulence factors have been discussed in detail. The high degree of sequence and functional conservation among dispartate virulence factors has been summarized succinctly. In these chapters, some of the very interesting topics which provide newer leads for further research are: periplasmic space of gram negative bacteria as a potential target for antimicrobial agents, study of *E. coli* hemolysin in conjunction with lipopolysaccharide to unravel its exact role in pathogenesis and tyrosine phosphatase activity of the yops of pathogenic yersiniae and their homology to eukaryotic signal transduction proteins.

Part 5 – 'Virulence gene regulation' – discusses molecular genetic mechanisms underlying regulatory loops responsible for the expression and regulation of virulence genes. The complexity of these regulatory loops has been illustrated by the well-known *bvg* regulation in *Bordetella pertussis* (Chapter 27) and methylation-dependent and *Lrp* (*L*eucine *r*esponsive regulatory *p*rotein)-dependent fimbrial gene expression in *E. coli* (Chapter 28). In addition to these, two other systems, namely urease as a potent virulence factor (Chapter 29) and plas-

mid-mediated iron transport virulence system of *Vibrio anguillarum* (Chapter 30), have been discussed. The chapter on urease makes very interesting reading wherein genetic organization and regulation of urease genes of members of enterobacteriaceae (*Proteus mirabilis*, *Klebsiella aerogenes* and *E. coli*), gram positive bacteria (*Staphylococcus saprophyticus* and a thermophilic *Bacillus* sp.) and *Helicobacter pylori* have been discussed.

Part 6 – 'Other aspects of bacterial pathogenesis' – is an assemblage of assorted aspects of bacterial pathogenicity which arose during Falkow's work detailed in the preceding chapters. None of these topics have been dealt with in previous compilations on the subject. It is apparent that these aspects of bacterial pathogenicity would constitute important areas of research in the future.

The enzyme systems responsible for sucrose transport, sucrose hydrolysis and synthesis of extracellular and intracellular polymers have been viewed as pivotal virulence factors for the cariogenic *mutans* streptococci. Newer tools which may help elucidate molecular mechanisms underlying synthesis, regulation and interrelations of these enzymes have been discussed (Chapter 31).

In general, it is very well-known that several factors, both host and pathogen, determine the minimum inoculum required to produce a disease in susceptible hosts. In this regard genetic determinants of acid resistance of *Shigella* spp. have been discussed (Chapter 32). However, the thought-provoking question in this chapter is: Are there any specific factors which dictate infective dose of a pathogen?

A great deal of progress has been made in recent years in the field of mucosal vaccines. Various aspects of the subject discussed include – cholera toxin subunit B as a model for the development of nonliving mucosal vaccines, its use as an adjuvant for other mucosal vaccines and live mucosal vaccines as carriers of heterologous antigens (Chapter 33).

The discussion on phylogenetic diversity of microbial pathogens (Chapter 34) brings into focus the bias of *in vitro* growth enrichment methods towards bacteria isolated from the environment. To what extent are these methods relevant to microbial flora of mammals? The inadequacy of these methods has been illustrated by the identification of unculturable pathogens directly from host tissue using PCR-amplified 16S rRNA sequences.

Since this volume is meant to be a tribute to Stanley Falkow, all the contributors are his former students, postdoctoral fellows or collaborators. So only those aspects of bacterial pathogenicity which have been pursued in his laboratory are reviewed. Nevertheless, it encompasses not only a good range of microbial pathogens and virulence determinants but also incorporates certain aspects, especially in Part 6, hitherto not dealt with in any of the previous works on this subject.

The book has the characteristic stamp of the American Society for Microbiology, i.e. highly authoritative and immensely readable.

All in all, this volume should form an important part of any collection, whether individual or library, dealing with bacteriology, medical microbiology, bacterial pathogenesis or infectious diseases.

J. S. VIRDI

Department of Microbiology,
University of Delhi South Campus,
Benito Juarez Road,
New Delhi 110 021, India.

Life in Moving Fluids. Steven Vogel. Princeton University Press, 41, William Street, Princeton NJ 08540, USA. 1994. 2nd edn. 467 pp. Price: $ 49.50, £ 35.

Scientists of yesteryears were known as natural philosophers: people who wondered about the workings of Nature. They appear to us as free souls, knowing no boundaries of specializations, and venturing into any body of knowledge to seek an explanation to some riddle that caught their imagination. The book *Life in Moving Fluids* by Steven Vogel is a beautiful one, one that revives the memories of the natural philosopher. It is an account of a biologist stepping out of his customary confines and thinking about the effects of flow on life in moving fluids.

Fluid mechanics has become a subject highly mathematized with a quest for exact results. The advent of high-speed computers has only enhanced this trend, and the frontier is occupied by computations of creeping flows, direct numerical simulations of turbulent flows, etc. All this is not a criticism of what goes on in fluid mechanics, but one does wonder: Is there any room for natural philosophers in fluid mechanics? Vogel's answer is an emphatic yes.

The book is structured as follows: Each chapter begins with an introduction of a set of fluid-mechanical concepts, followed by the role they play in a wide variety of biological phenomena. The discussion is clear and simple, and the speculations are intriguing. The author also describes many simple experiments done by himself. The gentle humour which he employs in his exposition is a bonus.

The book begins with the notion of a fluid and that all important property – kinematic viscosity. Here we learn how viscosity plays an important role in the lives of antarctic mammals and birds, and affects the morphology of some microcrustacean species. Then follow the very important concepts of kinematics: conservation of mass and streamlines. Based on these simple ideas one can calculate the number of capillaries in our body, even without needing the proverbial back of an envelope. The next topics are pressure and momentum. A clean derivation of Bernoulli's equation, and nice examples of flow measurements are presented. The ideas of pressure distribution and lift forces follow naturally, giving rise to beautiful examples of how these principles are 'used' by organisms to generate secondary flows, e.g. by termites to ventilate their mounds. All this leads up to a discussion of Newton's second law. It is nicely tied up with jet propulsion. It is really incredible that many living beings move by this principle, and belong to the 'jet set'!

The next few chapters drag 'drag' into focus. The new player is the Reynolds number. Here the origin of the drag force, the vagaries of the drag coefficients, separation of flow, streamlining of shapes, and scale modelling are explained in a very clear manner. Some of the author's own clean and simple experiments on scale modelling are presented. Application of the principles of drag to biological systems leads to an immediate complication since the shapes of living things are flexible and hence can interactively change with flow. Thus, there can be no discussion of 'dragginess' of a shape. All these ideas are discussed in the context of trees, microalgae, motile animals, etc.,

giving scope for some interesting questions involving evolution of shapes.

The next subject is velocity gradients, and boundary layers are introduced in a simple way laying emphasis on the geometries encountered in biological systems. Vogel captures the essence of boundary layer by saying that 'it is a place to hide from drag'. But, of course, hideouts have disadvantages too, since then there is a barrier to exchange of matter and energy. Consequently, there arise a whole lot of optimal existential problems. Thus, 'there is a price in drag to be paid versus paying in the harder currency of lift' due to their asymmetric shapes, cost of drag to be balanced against being away from the source of food, and penalty to be paid in terms of dispersal of spores for saving on drag, etc.

Then come the important concepts of vorticity, circulation and lift. Common mechanisms by which vorticity is generated are discussed followed by the topics of gliding, soaring, flying, and swimming.

The internal flows occupy the attention of the author in the next few chapters. As usual, the prototypical pipe flow is discussed. Here again the balance between power required to maintain the flow versus the decreased transport rates from wall in the absence of convection is brought to focus. Another problem faced by big organisms is the distribution of nutrient-carrying fluids to various corners of the body. This gives rise to Murray's law with its wide-ranging validity. This is fairly important given that the fraction of the metabolic rate spent on keeping the flow is considerable.

The next topic covers the low Reynolds number external flows. Here the flow is slow and the transfer rates are also low. Diffusion is the source of food, and one would have to work very hard to move in order to enhance the transport rates. Though motility is available, one is perhaps better off waiting, a fact humorously explained by the notion of 'a casual cow who, after eating, just waits for the local grass to regrow'. No discussion of creeping flows is complete without a discussion of terminal velocities. Vogel points out its relevance to survival, by citing the various structures developed by airborne particles – pollen grains, etc. – in order to maximize the dispersal distance, which has a direct bearing on fitness. The long rangedness of low Reynolds number flows is beautifully brought out by showing its relevance to information transfer by flow disturbances in the environment. The creepy games of preys and predators are subject to this feature of the creeping flows. The book ends with a discussion of unsteady flows and effects of interfaces.

Vogel intended this book to be treated, in part, as an argument that the unavoidable imperatives of mechanics of fluids underlie the biological design. He succeeds eminently in his presentation.

Here the interview of G. I. Taylor by G. K. Batchelor comes to my mind (*J. Fluid Mech.*, 1975, **70**, 625). In response to a query about the important areas in fluid mechanics, Taylor mentions the problem 'Why does a tree grow upward?'. What role did Taylor see in this problem for fluid mechanics, and, if he were alive, what would he have done regarding this or other such issues? Great minds are difficult to comprehend but it is easy enough to see that he was wondering about nature and was seeking the role of flow in nature. Taylor would certainly have given beautiful and simple resolutions of some of the issues raised in this book by Vogel.

I recommend Vogel's book to all interested in fluid mechanics. Even conventional courses on fluid mechanics can be enriched and made interesting by choosing elements from it.

K. S. GANDHI

Department of Chemical Engineering,
Indian Institute of Science,
Bangalore 560 012, India.

Annual Review of Neurosciences 1995. W. Maxwell Cowan. Annual Reviews Inc., 4139 El Camino Way, Palo Alto, P.B. No. 10139, California 94303-0139, USA. Price: USA $ 47, elsewhere $ 52. 607 pp.

This issue of the *Annual Review of Neurosciences* covers a wide range of interests. The first article by Phillipe Soriano is on the use of embryonic stem cells for making transgenic mice. This powerful technique has been used to develop targeted mouse mutants for specific neuronal receptors. The article deals primarily with the technical advances made in this field, which have been considerable. However, for those of us interested in learning more about the contribution of targeted mice mutants to the understanding of neurobiology, the article serves more as a catalogue. A perspective on fresh neuroscientific insights obtained as a consequence of this method would have been welcome. A number of articles review the recent advances made in the understanding of various human neurobiological diseases. Of these, the article by Suter and Snipes on peripheral neuropathies, has integrated nicely knowledge from basic molecular research on the components of the myelin sheath, with information from genetic and clinical studies. The article by Warren and Ashley Jr on the fragile X syndrome sheds important light on the role of triplet repeats in causing genetic mutations. On reading this review it appears likely that the triplet repeat expansion mechanism may serve as a cause of many, as yet poorly understood, neurological disorders. A suggested role for genomic imprinting during this process makes this mechanism all the more intriguing. The review article on neurobiology of infantile autism (Ciaranello and Ciaranello) highlights the lack of understanding regarding the cause(s) of this human disease. It appears to have a genetic basis and may in some cases be associated with other neurological disorders.

A number of articles in this review have tried to put together recent experiments in understanding the development of the vertebrate central nervous system (CNS). Hatten and Heinz discuss the general principles that have emerged from studies of the development of the cerebellar cortex. This particular brain tissue serves as a good model system since it is relatively less complex and its structure is well-understood. Analysis of its development has been further aided by the presence of mutants in which parts of the cerebellar cortex do not develop properly. Colamarino and Tessier-Lavigne have reviewed the vast array of information implicating molecules in the floor plate as cues for axon guidance in the developing CNS. Many of these molecules still remain to be identified while others such as Netrin-1 and aminin are being studied extensively. Another interesting molecule is agrin, which has been shown to play an important role in the formation of the neuromuscular synapse. Bowe and

Fallon in a concise review have put together much of the new data about agrin and the proposed mechanism by which it causes clustering of acetylcholine receptors at the postsynaptic junction. Current evidence also points strongly to a role for this molecule in synapse formation in the CNS.

The article by Ranganathan, Malicki and Zuker has made a strong case for studying sensory transduction processes in *Drosophila*. The combined genetic and molecular approach, possible in this organism, has helped identify genes required at almost every step of the visual transduction pathway. As stated by the authors, while there is a great deal of similarity in the overall strategy of visual transduction between vertebrates and invertebrates, much of the underlying molecular machinery is different. Another field in which *Drosophila* molecular genetics has had an important contribution is that of circadian rhythms. In his review on the molecular neurobiology of circadian rhythms in mammals, Takahashi tells us how attempts are being made to understand the molecular basis of this behaviour using both physiological and genetic experiments. The review argues for a molecular genetic approach in mouse similar to that taken by Konopka and Benzer twenty years ago in *Drosophila*.

This issue also provides a good introduction to some new and exciting areas in neuroscience. The field of integrative processing is one such topic reviewed by Knudsen and Brainard. It looks at the two better understood stimuli – vision and audition – and their integration in the brain. Principles that emerge from the study of these two stimuli may be useful in studying other integrative processes. Another interesting new area where research is still fairly preliminary is that of plasticity in the adult sensory cortex. This has been reviewed by Weinberger. It is still too early to say what the level of plasticity is in the adult sensory cortex and the suggestion from this review is that both plastic and nonplastic populations of cells may exist. Their relative roles in adult cortical functioning await further research.

GAITI HASAN

National Centre for Biological Sciences, TIFR Centre, IISC Campus, P.O. Box 1234, Bangalore 560 012, India.

Textbook of Molecular Biology. K. Sivarama Sastry, G. Padmanaban and C. Subramanyam. Macmillan India Ltd., 2/10 Ansari Road, New Delhi 110 012. pp. i–xiii + 498. Price: Rs. 168.

Molecular biology is all pervasive in life sciences today. Although it started as a rather narrow field, the commonality of biological information and its application in living systems caused this discipline to break all barriers of specialized fields in biology so that now there is no area in biology which has not been influenced by molecular biology and which has not influenced molecular biology. Therefore, a study of molecular biology is imperative for any student of biology. Since molecular biology spans a very wide canvas, a book on this topic needs to cover a much wider area than most other subjects in biology. In this context, the book by Sastry, Padmanaban and Subramanyam is a welcome addition to the many books available in this general area. However, compared to most of the other books published in India in this area, the present book is more comprehensive and better organized. The authors have done a remarkable job in putting so many different areas together in a relatively simple narrative style which makes it easy reading for young students, to whom it is primarily addressed. Relevant topics from biochemistry, genetics, cell biology, cytology, immunology, cancer biology, genetic engineering, etc., have been amalgamated to produce this book. With so many different topics to be covered, the order of their appearance in the book can be a matter of individual choice, but as the authors point out in their preface to the book, each chapter can also be read as stand-alone since extensive cross-references to different topics help the reader navigate through related topics. Inclusion of references to more extensive books as also to articles on specific topics is very helpful.

With a constraint on the overall volume of the book (presumably to keep the price within reasonable limits), the book provides, in most cases, only the basal information on any given topic without going in for a detailed exposition. This limits the utility of the book for students at different levels of studies and for those interested in certain topics in more detail. Nevertheless, the book can be used by most students and teachers either as a primer for details in other more extensive treatises available or as a summary after having read a given topic in more detail elsewhere. In either case, the book will be useful for the student community.

This book is fairly strong in the biochemical aspects of molecular biology, as may be expected. However, factual and/or conceptual errors exist in the book, particularly in topics that relate to genetics, cell biology and developmental genetics, etc. (these are the ones that I read more carefully due to my own bias). Rather than making an exhaustive listing of these, a few examples may be noted. The $2N$ of *Drosophila melanogaster* is given as 4 instead of 8 (p. 90). The chapter on genome organization includes a brief discussion on X-inactivation in mammals (p. 111); unfortunately, however, this discussion does not make it clear that X-inactivation occurs only in somatic cells of female mammals (also the term 'lyonization' is after 'Mary Lyon' rather than 'Mary Lyons').

While discussing recombination during meiosis, the authors have also considered sister chromatid exchanges and imply that sister chromatid exchanges are central to recombination (p. 353, last para); the description gives an impression as if sister chromatid exchanges and chiasma points in meiosis are similar. While sister chromatid exchanges in mitosis and the homologous recombination during meiosis (which correlates with the cytologically visible chiasma, different from sister chromatid exchange) may share some of the enzymatic pathways, the two are distinctly separate phenomena with very different genetic consequences. In the absence of a more detailed discussion, the reader would remain misinformed on these issues.

Consideration of the genetic control of development in *Drosophila* on pp. 420–424 is rather incomplete and confusing. The classification of segmentation genes into three categories (the gap, pair rule and segment polarity genes) was proposed by Nusslein-Volhard and Wieschaus in 1980 rather than by Gehring, as the book seems to suggest (p. 421). The *fushi-tarazu* gene is not a part of the *Ant-C* complex: although it is physically included in the *Ant-C* complex DNA region, it acts independently as a pair rule gene rather than as a member of the *Ant-C* homeotic gene complex. Another common

misconception, that seems to have affected the description in this book as well, is about the term homeotic genes and homeo-box: all genes that carry the homeo-box sequence are not homeotic genes. I think the elucidation of specific details of the cascade of gene activity (beginning with the maternally acting genes and followed by the gap, pair rule, segment polarity and homeotic genes in that order) that defines the embryonic segmentation and fate determination in *Drosophila* and other higher animals has been one of the exciting hallmark achievements of the combined use of molecular biology, genetics and developmental biology. This chapter does not convey the specific information now available on this cascade of embryonic gene activity. The presentation in the book gives impression of a disjointed piecemeal information on certain genes and thus fails to carry the excitement to readers.

The chapter on genetic engineering is essentially confined to some of the basic recombinant DNA techniques (random-primed labelling deserved a place here). To justify the topic fully, it could have included examples where genetic engineering techniques have actually been employed as also the potential uses and hazards of these technologies.

While the book is generally well illustrated, the figures could be more attractive and self-sufficient: this would allow a better comprehension to the reader. This text book has made a good beginning. It is to be expected that its subsequent editions will take care of the shortcomings of the type noted above so that it evolves into a popular and authoritative textbook that is badly needed by our student community.

S. C. LAKHOTIA

Cytogenetics Laboratory,
Department of Zoology,
Banaras Hindu University,
Varanasi 221 005, India.

Corrections

Bose–Einstein condensation in a dilute atomic vapour

N. Kumar
[*Curr. Sci.*, 1995, **69**, 492–493]

Read '... the condensate size is "greater" than ...' instead of '... "less" than ...' in column 2, line 2 from the bottom of the article.

Evaluation of efficacy of bricks as geochemical monitor of atmospheric heavy metal pollution

Rohit Shrivastav, Sanjay Kumar Mathur, Shobhit Shrivastav, M. M. Shrivastav, Sahab Das and Satya Prakash
[*Curr. Sci.*, 1995, **69**, 272–275]

The mean concentration of metals in Table 2 and Table 3 should read mg per 100 g instead of mg kg^{-1}.

Edited and published by Prof. S. Ramaseshan and Prof. P. Balaram, Current Science Association, Bangalore 560 080.
Typeset by R D Typesetters, Bangalore 560 010. Printed by Lotus Printers, Bangalore (Ph: 3380167).

CURRENT SCIENCE

A fortnightly journal of research

Editors: S. Ramaseshan and P. Balaram

Editorial Board

S. Arunachalam, Central Electrochemical Research Institute, Karaikudi 623 006
Asis Datta, School of Life Sciences, Jawaharlal Nehru University, New Delhi 110 067
D. Balasubramanian, Centre for Cellular and Molecular Biology, Hyderabad 500 007
J. Chandrasekhar, Department of Organic Chemistry, Indian Institute of Science, Bangalore 560 012
R. Gadagkar, Centre for Ecological Sciences, Indian Institute of Science, Bangalore 560 012
K. N. Ganeshaiah, Department of Plant Genetics and Breeding, University of Agricultural Sciences, Bangalore 560 065
R. V. Hosur, Tata Institute of Fundamental Research, Bombay 400 005
N. V. Joshi, Centre for Ecological Sciences, Indian Institute of Science, Bangalore 560 012
Jyotirmoy Das, Indian Institute of Chemical Biology, Calcutta 700 032
C. C. Kartha, Division of Cellular and Molecular Cardiology, Sree Chitra Tirunal Institute for Medical Sciences and Technology, Thiruvananthapuram 695 011
S. S. Krishnamurthy, Inorganic and Physical Chemistry Department, Indian Institute of Science, Bangalore 560 012
S. C. Lakhotia, Department of Zoology, Banaras Hindu University, Varanasi 221 005
K. Muralidhar, Department of Zoology, Delhi University, Delhi 110 007
R. Narasimhan, CMC Limited, Bangalore 560 001
K. Neelakantan, ANURAG (Advanced Numerical Research and Analysis Group), Hyderabad 500 258
R. Nityananda, Raman Research Institute, Bangalore 560 080
T. Padmanabhan, Inter-University Centre for Astronomy and Astrophysics, Pune 411 007
B. R. Pai, Propulsion Division, National Aerospace Laboratories, Bangalore 560 017
T. J. Pandian, School of Biological Sciences, Madurai Kamaraj University, Madurai 625 021
Y. S. Rajan, Department of Science and Technology, New Delhi 110 016
M. Ramakrishnan, Geological Survey of India, Calcutta 700 016
A. V. Ramani, T. T. K. Pharma Ltd., Bangalore 560 025
S. Ramaswamy, Department of Physics, Indian Institute of Science, Bangalore 560 012
A. K. Raychaudhuri, Department of Physics, Indian Institute of Science, Bangalore 560 012
S. R. Shetye, Physical Oceanography Division, National Institute of Oceanography, Dona Paula, Goa 403 004
V. Siddhartha, 51 Bharati Nagar, New Delhi 110 003
V. Srinivas, School of Mathematics, Tata Institute of Fundamental Research, Bombay 400 005
G. Srinivasan, Raman Research Institute, Bangalore 560 080
R. Uma Shaanker, Department of Crop Physiology, University of Agricultural Sciences, Bangalore 560 065
K. S. Valdiya, Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore 560 064
G. Venkataraman, Department of Physics, Sri Sathya Sai Institute of Higher Learning, Prasanthi Nilayam 515 134
M. Vidyasagar, Centre for Artificial Intelligence and Robotics, Bangalore 560 001
K. VijayRaghavan, TIFR Centre, Indian Institute of Science Campus, Bangalore 560 012

Current Science, founded in 1932, is published by the Current Science Association in collaboration with the Indian Academy of Sciences. The journal is also intended as a medium for communication and discussion of important issues that concern science and scientific activity. All articles published in *Current Science*, especially editorials, opinions and commentaries, letters and book reviews, are deemed to reflect the individual views of the authors and not the official points of view, either of the Current Science Association or of the Indian Academy of Sciences.

Working Committee, Current Science Association

P. Balaram, Indian Institute of Science, Bangalore
R. Chidambaram, Department of Atomic Energy, Bombay
V. K. Gaur, CSIR Centre for Mathematical Modelling and Computer Simulation, Bangalore
K. Kasturirangan, ISRO Satellite Centre, Bangalore
P. K. Kaw, Institute for Plasma Research, Gandhinagar
N. Kumar, Raman Research Institute, Bangalore (*Vice-President*)
N. V. Madhusudana, Raman Research Institute, Bangalore (*Secretary*)
A. Mani, Field Research Unit, Indian Institute of Tropical Meteorology, Bangalore (*Past-President*)
R. A. Mashelkar, Council of Scientific and Industrial Research, New Delhi
R. Narasimha, Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore
G. Padmanaban, Indian Institute of Science, Bangalore (*Vice-President*)
R. Ramachandran, The Institute of Mathematical Sciences, Madras
P. Rama Rao, Defence Research and Development Organization, New Delhi
S. Ramaseshan, Raman Research Institute, Bangalore (*Treasurer*)
M. S. Valiathan, Manipal Academy of Higher Education, Manipal (*President*)

Editorial Office

Editorial staff

Chandrika Ramesh, G. Madhavan, T. D. Mahabaleswara, C. S. Ravi Kumar, M. S. Venugopal

Circulation and Accounts

Peter Jayaraj, Ranjini Mohan, B. Sethumani, Shanthi Bhaskar, B. K. Shivaramiah

Annual subscription 1995

	Institutional	Personal
India	Rs 350	Rs 100
All other countries (AIR MAIL)	US $175	US $50

Single copies other than special issues: Rs 35/US $15

Editorial Office: CURRENT SCIENCE
C. V. Raman Avenue
P.B. 8001
Bangalore 560 080, India
Telephone: 91-80-3342310
Telex 845-2178 ACAD IN
Fax: 91-80-3346094

Copyright © 1995, Indian Academy of Sciences

INFORMATION FOR CONTRIBUTORS

GENERAL

All manuscripts should be addressed to the Editor, *Current Science*, P. B. No. 8001, C. V. Raman Avenue, Bangalore 560 080. Submission of an article will be held to imply that it has not been previously published and is not under consideration for publication elsewhere; and further, that if accepted, it will not be published elsewhere. *Three copies of contributions of all categories* are required, with a letter of transmittal giving (i) names and complete addresses of the authors and (ii) title of the contribution and the category in which it is submitted (see below).

Current Science is a multidisciplinary journal and therefore research and review papers of general significance that are written clearly and well organized will be given preference. All papers will be first assessed by a Reviewing Editor. Papers found unsuitable in terms of the overall requirements of the journal will be returned to the authors. The others will be sent for detailed review. *Both solicited and unsolicited material will be reviewed.* Authors of these papers will be notified of acceptance, rejection, or need for revision of the paper. Returned papers cannot be resubmitted.

Illustrations and other material to be reproduced from other publications must be properly credited; it is the authors' responsibility to obtain permission for reproduction (copies of letters of permission should be sent).

CATEGORIES OF MANUSCRIPT

General articles (not exceeding 5000 words) discuss current trends in research in a field that will be of interest to readers outside the field; interdisciplinary topics: science policy and administration; or some aspect of the application of science and technology to human needs or the impact of science and technology on society/ecosystems/life. They should include a summary not exceeding 100 words, introductory paragraph(s), brief subheads at appropriate places to point to what follows, illustrations that will help a general reader, and references.

Review articles (not exceeding 5000 words) are expected to survey and discuss recent developments in a field. They should be well focused and organized, and avoid a general, 'textbook' style.

Research articles (not exceeding 4000 words) should report research results of fairly major significance. They should include an abstract not exceeding 100 words, introductory paragraph(s), and brief subheads.

Research communications (not exceeding 2000 words) should contain important findings that are novel and of fairly broad interest. They should include a brief abstract and an introductory paragraph. Text should not be broken up under subheads.

Correspondence includes letters, not exceeding 500 words, that are of general interest to scientists. All letters cannot be published.

Scientific correspondence contains technical comments, including those on articles or communications published in *Current Science* within the previous six months. Letters may be reviewed and edited.

Research news articles are intended to inform nonspecialists about recently published advances or important findings discussed at a meeting. Authors should also send a copy of the paper(s) on which the article is based. Meeting reports should avoid merely listing brief accounts of topics discussed, and must convey to readers the significance of an important advance.

Research accounts articles are intended to be personalized reviews of research from the authors' own laboratory, based on a body of published work. The articles must provide appropriate background to the area in a concise introduction, which should also serve to place the author's work in proper perspective. Articles will normally not exceed 8 to 10 printed pages.

Opinion articles present views on issues related to science and scientific activity. **Commentary** articles should contain expository notes on issues related to science and scientific activity.

Book reviews. Unsolicited reviews will also be considered. Reviews that merely 'list' brief descriptions of the contents cannot be published. Reviews should have 'context' and convey some information about the subject of the book.

Historical commentary and notes inform readers about interesting aspects of personalities or institutions of science or about watershed events in the history/development of science; most are expected to relate to India. Illustrations are welcome. Brief items will also be considered.

MANUSCRIPT PREPARATION

Manuscripts should be typed double-spaced on one side of white bond paper (21 × 28 cm). The pages should be numbered consecutively, starting with the title page and through the text, reference list, tables and figure legends. The **title** should be brief, specific and amenable to indexing. Not more than five **keywords** should be indicated separately; these should be chosen carefully and must not be phrases of several words. **Summary** and **abstract** should not have more than 100 words and should convey the main point of the paper, outline the results and conclusions, and explain the significance of the results.

Text. All papers should have a brief introduction. The text should be intelligible to readers in different disciplines and technical terms should be defined. Tables and figures should be referred to in numerical order. All **symbols** and **abbreviations** must be defined, and used only when absolutely necessary. Superscripts and subscripts and ambiguous characters should be clearly indicated. **Units of measure** should be metric or, preferably, SI. Methods should, as far as possible, be described briefly in appropriate table and figure legends.

Figures. In the case of line drawings one set of originals (without any lettering) is sufficient, with two copies containing lettering. In the case of photographs good prints are required with each copy of the manuscript; photocopies are not acceptable. Line drawings should be roughly twice the final printed size. The correct orientation should be indicated if not clear.

Photomicrographs and other photographs that require it must have a scale bar, which should be defined clearly in the legend. Primary data should be submitted as far as possible (e.g. actual photographs of electrophoretic gels rather than idealized diagrams).

References should be numbered in the order in which they appear, first through the text and then through table and figure legends. The following are examples of ways of writing references.

1. Mukundan, T. and Kishore, K., *Curr. Sci.*, 1991, **60**, 355–362.
2. Constantine, G., in *Biology of Bats* (ed. Wimsatt, W. A.), Academic Press, New York, 1970, vol. 1, pp. 319–322.

Acknowledgements should be brief. Footnotes are not allowed except to identify the corresponding author if not the first.

Cover photographs. Good photographs (colour or black and white) that pertain to a submitted paper will be considered for use on the cover. Good *prints* and a legend should be submitted with the manuscript. In the case of a colour picture, a transparency will be required for printing if accepted.

PROOFS AND PUBLICATION

Two sets of galley proofs are sent to the corresponding author. A reprint order form accompanies the proofs.

Founded: 1932

ISSN: 0011-3891

CURRENT SCIENCE

Volume 69 Number 8

25 October 1995

631 In this issue

CORRESPONDENCE

633 **Ph D students, research supervisors and university research: Some observations**
S. John

SCIENTIFIC CORRESPONDENCE

634 **Recent finds of ancient gold mining sites in south Uttar Pradesh**
R. V. Vyas, H. R. Patel, B. A. Patel and D. J. Patel

635 **Parasitism of muscardine fungi on *Meloidogyne incognita* eggs**
G. N. Dwivedi, R. C. Tyagi, A. K. Tripathi and M. L. Yadav

COMMENTARY

637 **Intellectual property rights and biological resources: Specifying geographical origins and prior knowledge of uses**
Madhav Gadgil and Preston Devasia

GENERAL ARTICLE

640 **Developments and trends in inductively coupled plasma mass spectrometry and its influence on the recent advances in trace element analysis**
V. Balaram

649 **Climate change and agriculture – An Indian perspective**
Sulochana Gadgil

REVIEW ARTICLES

659 **Opioid peptides in invertebrates: Localization, distribution and possible functional roles**
R. Nagabhushanam, R. Sarojini, P. S. Reddy, M. Devi and M. Fingerman

672 **Carbon dioxide and nitrous oxide in the north Indian Ocean**
M. Dileep Kumar, S. W. A. Naqvi, D. A. Jayakumar, M. D. George, P. V. Narvekar and S. N. de Souza

RESEARCH ARTICLES

678 **Internal waves – A novel measurement technique**
Rao Tatavarti, P. N. Ananth, K. Rajasree, V. Vidyalal, P. Radhakrishnan, V. P. N. Nampoori and C. P. G. Vallabhan

685 **Habitat suitability analysis of *Nemorhaedus goral* – A remote sensing and geographic information system approach**
P. S. Roy, Shirish A. Ravan, N. Rajadnya, K. K. Das, Abhineet Jain and Sarnam Singh

RESEARCH COMMUNICATIONS

691 **Carbon and oxygen isotopic records of planktonic and benthic foraminifera from a new deep-sea core of the northeast Indian Ocean**
S. Masood Ahmad

695 **Resistance to *Bacillus sphaericus* in *Culex quinquefasciatus* Say 1823**
T. Adak, P. K. Mittal, K. Raghavendra, Sarala K. Subbarao, M. A. Ansari and V. P. Sharma

698 **Lead impairs hepatic type I-5′-monodeiodinase activity and thyroid function in cockerels**
S. S. Chaurasia, P. Gupta, A. Kar and P. K. Maiti

BOOK REVIEWS

701 **Speaking Minds: Interviews with Twenty Eminent Cognitive Scientists,** *reviewed by* R. Narasimhan

702 **Development of Ideas in Physics,** *reviewed by* Arvind Kumar

702 **A Modern Introduction to Ancient Indian Mathematics; Indian Mathematics and Astronomy,** *reviewed by* C. S. Yogananda

703 **Leks,** *reviewed* by R. J. Ranjit Daniels

HISTORICAL COMMENTARY AND NOTES

705 **Samanta Chandra Sekhar: Life and work**
P. C. Naik and L. Satpathy



COVER. *Tinospora cordifolia* (Willd) Miers ex Hook. F. & Thomas. This climbing shrub distributed widely in India and Sri Lanka is used in medicated wines, aqueous extracts and starch extracts as an immunomodulatory drug against fever and rheumatic diseases in the Ayurvedic system. There are extensive opportunities for exploiting such plants and traditional knowledge of their use in the modern biotechnology industry. See page 637 for a proposal to facilitate equitable sharing of benefits of such applications. [Photo courtesy: Foundation for Revitalization of Local Health Traditions, Bangalore]

Indexed in CURRENT CONTENTS/GEOBASE/CHEMICAL ABSTRACTS

In this issue

Opioid peptides: Two decades on

Two decades have passed since Hughes, Kosterlitz and their colleagues reported the discovery of the enkephalins from porcine brain (*Nature*, 1975, **258**, 577). These pentapeptides were the first endogeneous opioids to be discovered providing a new class of ligands to explore the physiology and pharmacology of an ever-increasing number of opiate receptors in the brain. Classically, morphine – an opium alkaloid, was the best known of the opiate receptor ligands. In the early years of opioid peptide research, much attention was focused on the possibility of developing non-addictive analgesics, a goal which proved unreachable despite enormous scientific effort. The discovery of endogeneous opioids, however, proved to be extraordinarily influential in directing attention towards neuropeptides, their role in pain perception and modulation of behaviour. Close connections were established between the brain and the gut, a fact that is not generally appreciated. While initial discoveries were made in a variety of vertebrates, very soon opioid peptides were found to be widely distributed among invertebrates, providing comparative endocrinologists with a fertile area of research. In a comprehensive review (**page 659**), Nagabhushanam *et al.* survey the occurrence and biological role of peptide opioids in invertebrates. A vast body of literature implicates these molecules in a diverse range of biological phenomena, ranging from behaviour modulation and thermoregulation to reproduction. This spectacular range of molecules and functions could hardly have been anticipated when the first reports on enkephalins appeared twenty years ago.

P. Balaram

Of remote sensing, GIS, and the mountain goat

Wildlife biology has been one of those subjects where even the simplest questions are maddeningly difficult to answer (e.g. how many tigers are there in Nagarhole?), and the considered opinion of an expert practitioner is much more valuable than the results obtained from the most sophisticated mathematical models. Recent years, however, have seen a welcome trend of a close co-operation between the two; the availability of satellite imagery at increasingly finer scale and of powerful GIS (Geographical Information System) hardware/software has improved the situation even more. The article on **page 685** of this issue illustrates a particularly attractive contribution – of estimating the proportion of the area of the Rajaji National Park, where a population of *Nemorhaedus goral*, an endangered species of the mountain goat, can thrive.

Situated in the Dhaulkhand range in the Shivalik mountains of Uttar Pradesh, the Rajaji National Park covers an area of about 150 square kilometers. Contrary to the imagined picture of vast stretches of thick forest, the park is a rich and complex mosaic of dry deciduous forest patches, roads, human habitation, plantations, water bodies and the like. Not all of these are equally conducive for maintaining the population of mountain goats, however – they prefer areas with steep, grassy slopes and near water bodies, and tend to avoid locations close to human habitation, and those with thick shrubs.

Starting with satellite imagery down to the resolution of 36 m × 36 m, the authors were able to prepare detailed maps (2 cm = 1 km) describing the different vegetation types in the national park. To this was added the information on the location of habitations, slopes, waterbodies, etc. GIS software was then used to assign a numerical value (of index of suitability to gorals, based on what the gorals like and dislike) to each point in the map. Further analysis then showed about 35% area to be highly suitable to gorals, and 20% unsuitable to them. Clearly, such quantitative estimates would be very valuable to foresters and administrators of national parks for efficient management of wildlife.

A particularly heartening aspect of this article is the close collaboration of the different institutions – Remote Sensing Institute, the Forest Department and WWF – and one can now look forward to more of such interdisciplinary contributions based on the study of various national parks and other endangered species.

N. V. Joshi

Indo-Israel Joint Programme of Cooperation

Department of Science & Technology (DST) invites applications from Indian scientists for consideration for the exchange visits to Israel in the following categories:

(a) Senior Scientists	:	For a period of two weeks.
(b) Short-term exchange	:	For scholars with a Ph D degree on an invitation from Israeli scientists, for a period of 3–4 months up to the age of 40 years.
(c) Post-doctorate	:	For one academic year (10 months) for scholars with a Ph D degree awarded not more than two years prior to the application.

Air fare to the visiting Indian scientists will be provided by the Indian side and their local hospitality will be provided by the host country.

The priority areas presently identified for the exchange visit are Advanced materials, Microelectronics, Biotechnology, Information technology and Electro-optics. Excellet applicants in other areas of material sciences and engineering will also be considered.

The scientists interested in exchange programme may send their application duly forwarded by their institution along with a two-page pre-proposal (brief justification, specific area and duration) and a brief write up on the infrastructure available at the Institution of the application.

The application may be submitted to the following address:

Indo-Israel Joint Programme of Co-operation
International Division
Department of Science & Technology
Technology Bhavan
New Mehrauli Road
New Delhi 110 016

Last date for receipt of the application for the year 1995–96 is 30th November 1995.

CURRENT SCIENCE

Volume 69 Number 8 25 October 1995

CORRESPONDENCE

Ph D students, research supervisors and university research: Some observations

V. Sitaramam in his opinion 'Dynamics of the psychology of Ph D students and the question of what to do with them' (*Curr. Sci.*, 1995, **68**, 779–783) has considered only a part of the university research scenario. It is essential to look at the system as a whole. The foregoing account is based more on observations than useful statistics (which anyway is difficult to come by).

It is true that most of the research in the country (good, bad and *not so beautiful*) is carried out by research students. But they do it under the 'guidance' of their guides (research supervisors). Hence it is necessary to have a look at both these groups, Ph D students and the research supervisors, and the role their interaction plays in university research.

Students opting for Ph D can be grouped into two broad categories: (i) Those who have passed the UGC/UGC-CSIR NET (with a minimum assured standard) and (ii) Those who have not passed the same (nothing can be said about this category which may even consist of some persons having no academic merit, whatsoever).

The first category consists of two types of students: (i) those opting for Ph D because of their interest in the subject and interest in research, and (ii) those opting for a Ph D because of unemployment. The latter group is interested only in the money (the stipend), and considers Ph D programmes as employment. Students belonging to all categories get accepted for Ph D in our universities. This is because there are more unwritten laws belonging to the same category as Parkinson's fifth (?) law, which govern our university system.

One of these laws states '*Students opting for Ph D programmes can be graded on a scale of 0–10 on the basis of their intelligence/ignorance and competence/incompetence, and for every student there is a guide of the required level of competence/incompetence*'. This great miracle is possible because of another law which states '*When the authorities are interested, there is no dearth of subject experts of required degree of compliance who are ready to doctor the premature birth of anyone as a university teacher*' (a modified version of this is applicable in promotions). The Parkinson's fifth (?) law quoted by Sitaramam obviously is an extension of these two laws.

The trend of those scoring very high marks in 10th and 12th level opting for professional courses has been '*in*' for a long time. Hence the present generation of university teachers (research supervisors) is also drawn from the layers lying below the '*top-most*' layer.

The research supervisors (especially in the university set-up) can be grouped into three categories on the basis of their selection.

1. *Selection (purely) on the basis of merit.* This group consists of the best students returning from universities abroad, national institutions, from IITs, from the same university or other Indian universities. They have excellent academic records and have worked on good topics, with good guides, in good institutions. They have a good exposure, and have a deep interest in research (in specific areas). They co-operate well with the students, work hard, and publish frequently in good journals.

2. *Selection due to inbreeding.* In case of this group there is some compromise on the quality. They are better students from the same university, mostly students of those in authority. They also have a reasonably good exposure, are in general interested in research and helpful to the students.

3. *Selection due to favouritism.* In case of this group there is a heavy compromise on the quality. They have average or below average academic records. Mostly close relatives of those in authority, politicians or their friends, they have somehow managed a Ph D, have little exposure, are not much interested in research but do research because of compulsion (others are doing it, some research is necessary to justify the presence in the university environment and for promotions).

The fate of the Ph D student depends much on the guide he/she selects.

Fresh M Scs have enthusiasm. They want to do research but most of them do not know what exactly to do. They are in a state of *infatuation* with science. It is up to the guide to transform this *infatuation* into a true love for science and his/her initial interaction with science as a meaningful and long-lasting love affair. In other words, to understand that Ph D is a training programme and to guide and mould the student into a first rate researcher. Most of the research topics are suggested/selected by research guides. Fresh M Scs cannot be expected to come up with a suitable Ph D problem. No research is an end in itself. Most of them raise as many questions (if not more!) than they answer. Because of this, a guide having deep interest and

working constantly in any field will have enough problems for fresh students.

Now the question arises: 'What can be done to improve the situation?' Three things can be done:

1. The selection of the university teachers can be made more rigorous, if possible at the national level, than relying on local selection committees.

2. The UGC/UGC-CSIR NET can be modified with more emphasis on research aptitude, and

3. Universities can introduce entrance test for those who have not passed UGC/UGC-CSIR NET/GRE/GATE etc. for registering for a Ph D.

About the dynamics of the psychology of research students, I am aware of one. All students start their Ph D programme with an intention (hope!) of completing it in 3–4 years. Some of them who have teamed up with *not too enterprising* guides take very long (7–8 years). It takes 2–3 years for the student to realize that the guide is not very helpful. The student continues because he/she feels that in another 1–2 years the work can be completed. And it is better to spend a little more time than losing the efforts already put in for 2–3 years. After another year if the situation still remains the same, the student continues. As time passes, the time required to complete Ph D becomes increasingly a lower fraction of the time already spent. There are societal compulsions also (of relatives and friends scorning 'he/she is not intelligent enough to get a Ph D'). The end of this prolonged (perennial) suffering is when both the student and the guide get totally fed-up, and the guide decides to resort to the Parkinson's fifth (?) law. This leaves an incompetent guide with one more feather in his/her cap (of having produced one more Ph D) and a student who has lost all interest in research (possibly some interest in life also), overaged with a Ph D degree which at present does not have a ready market.

Finally, I would like to disagree with Sitaramam in his suggestion that we should plan our programmes on the basis of the average performers. In my opinion we must strive for excellence and excellence alone. Planning on the basis of average performer results in compromise. Compromises make more compromises necessary. And all these compromises spoil the system.

S. JOHN

Shibu Nivas, Sagar Park,
S. No. 46/9, Nagar Road,
Pune 411 014, India

SCIENTIFIC CORRESPONDENCE

Recent finds of ancient gold mining sites in south Uttar Pradesh

A prominent ancient gold mining site was recently discovered by the authors[1] in the Precambrian Shield region of Uttar Pradesh. It is located about 2.5 km west of Gurmura village in Sonbhadra district (Figure 1). The clusters of old workings are spread over an area about one km in strike length and over 600 m wide. It can be divided into northern, central and southern parts. The northern part shows seven elliptical to rectangular trenches measuring up to 30 m × 10 m × 10 m in size with an incline aimed to exploit easterly plunging ore body. The main hillock, locally referred to as Sona Pahari, forms the central part of the area of old workings. It has large trenches, the largest being about 60 m × 10 m × 10 m located on top of the Sona Pahari. The southern part extends over a length of about 500 m. This has two main trenches. The larger measures about 30 m × 5 m in plan. Both the trenches have inclines. Mishra[2] had speculated about the possibility of gold mining at Sona Pahari (golden hillock) on phonetic connotation and also on the basis of its location in the provenance of Ashokan rock edicts, which are strikingly common in gold-bearing regions of India.

Regionally, the early to middle Proterozoic rocks of southern UP, correlated with the Bijawar[3] Group, rest over a granitic basement, viz. the Dudhi Granitoid exposed to the south, and are overlain by late Proterozoic Vindhyan sediments exposed to the north (Figure 1). The area exposes a dominantly argillitic sequence with minor graywacke–arenite intercalations referred to as the Parsoi Formation[4] of Bijawar Group. The argillites show E–W to ENE–WSW-trending foliation with high dips generally towards south. The succession of metamorphosed chemogenic sediments comprises banded hematite quartzite/jasper, banded magnetite quartzite; argillite sequence consisting of phyllite–graywacke–arenite units and the associated volcanics consisting of a suite of basic to ultra-basic phases, agglomerates and tuffs. The Group compares well with the Precambrian Greenstone Belts[5]. The argillite sequence is intruded by a number of quartz veins with sympathetic trend emplaced probably synchronous to the emplacement of a granitic pulse, viz. the Jhirgadandi granite, in the adjacent area. The mineralized quartz veins are oxidized on the surface and exhibit pyrite, arsenopyrite and galena. Narrow zones of wall rock alteration, indicating carbonatization and kaolinization, signify hydrothermal action. The mineralization appears to be controlled by E–W to ENE–WSW-trending folds and the ancient miners seem to have excavated the richest portion of ore emplaced along the axial-plane shears of the doubly plunging folds with plunge around 35°.

The analytical results for the major trace elemental contents of 50 surface grab samples collected from the area indicate the presence of gold and silver along with other associated metallic elements. Forty-three per cent of samples analysed by Atomic Absorption unit at chemical laboratories, Geological Survey of India at Lucknow, Hyderabad

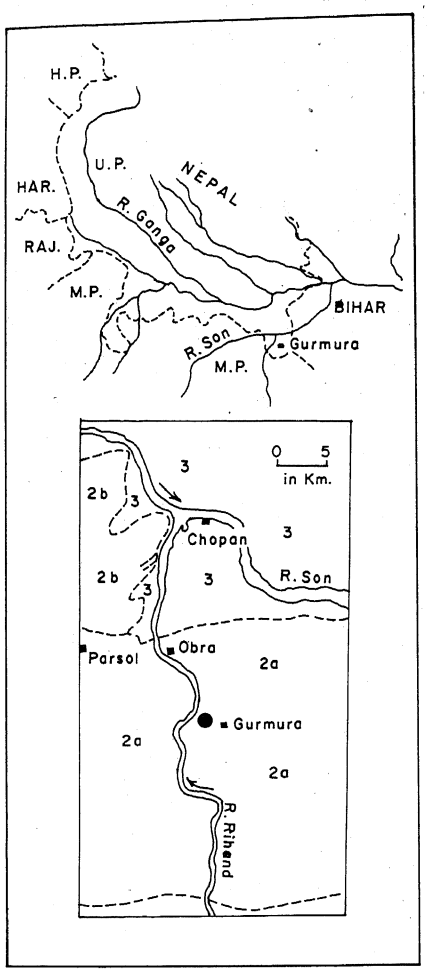


Figure 1. Geological map and location of ancient mining sites around Gurmura, Sonbhadra District, UP. Index: 3 – Vindhyan Super Group; 2b – Agori Formation; 2a – Parsoi Formation; 1 – Dudhi Granitoid; 0 – Location of old workings.

and Calcutta, indicated anomalous gold analysing from 0.10–2.80 ppm Au with 5–55 ppm Ag, 35–535 ppb Hg, up to 0.34% Cu and 0.12% Zn, 200–3000 ppm As, average 1.9% Pb and up to 20 ppm Cd (analysed in one sample). About 95% of anomalous values are from quartz vein samples and 5% are from argillite lithology. The anomaly distribution is shown in Table 1.

Table 1.

Lithology	Anomaly (%)	Au range (in ppm)	As range (in ppm)	Arithmatic mean for Au (in ppm)
Quartz vein	95	0.10–2.80	200–3000	0.83
Phyllite	5	0.16	n.d.	0.16

Preliminary limited analytical data indicate the presence of Au, Ag, Cd, Pb, Cu and Zn. The favourable geological set-up for mineralization and the associated extensive old workings make it a new prospect. Detailed geochemical sampling, presently in progress, is likely to reveal the exact nature of mineralization.

A noteworthy find includes the presence of a large number of crude stone crushers and mortars scattered throughout the area. The stone mortars have elliptical depressions measuring 20 cm × 15 cm in plan and are mainly of hard arenitic or doleritic rocks. The find suggests that grinding was an important stage in winning or concentrating the ore. The gold was extracted in a phased manner. Evidences of ancient gold mining and extraction in similar manner along with similar stone crushers and mortars have also been recorded from the contiguous western area around Sudda[6], Sonkorwa, Matehni and Gurhar Pahar[7], etc., situated in the Sonbhadra District, UP, and Sidhi district, MP. The absence of slag material, retorts, etc., associated with smelting of base metals, indicates that the base metals have not been extracted and all the mining activity was for gold only.

As regards the age of the ancient mining activity, the preponderance of stone crushers and mortars with occasional finds of iron chiesels, etc., point to ancient historical period. The age of gold mining is pre-Mauryan and might be pre 2400 BC.

1. Anonymous, *Geological Surv. India, Northern Region News*, 1993, **13**(3), 4.
2. Mishra, S. P., *Rec. Geol. Surv. India*, 1987, **113**(7), 92–98.
3. Auden, J. B., *Mem. Geol. Surv. India*, 1934, **63**(2).
4. Narain, Kedar, *Rec. Geol. Surv. India*, 1964, **89**(1), 7–8.
5. Chaubey, V. D. and Gupta, Arindam, *J. Geol. Soc. India*, 1990, **35**, 296–305.
6. Singh, M. and Khan, M. A., Report on petrochemical studies of igneous suite of rocks, Obra–Garda Greenstone Belt, Mirzapur district, UP, Geol. Surv. India, Progress Report, unpublished, 1985.
7. Lal, J. K., Sharma, D. P., Mehrotra, S. C. and Kannadasan, T., *Rec. Geol. Surv. India*, **122**(8), 151–152.

ACKNOWLEDGEMENT. We are grateful to Shri K. Krishnanunni and Ravi Shankar, Deputy Directors General, G. S. I., for suggestions.

G. N. DWIVEDI
R. C. TYAGI
A. K. TRIPATHI
M. L. YADAV

Geological Survey of India, Lucknow 226 020, India

Parasitism of muscardine fungi on *Meloidogyne incognita* eggs

Root-knot nematode, *Meloidogyne incognita* (Kofoid and White) Chitwood, is a serious pest in vegetable and pulse crops in India[1]. Several biocontrol agents have been reported on root-knot nematodes the world over. Some of the bioagents are found very useful in managing the nematode problem in agricultural crops[2]. Hence, in this context, two important entomopathogenic muscardine fungi, *Metarhizium anisopliae* (Metch.) Sorokin and *Beauveria brongniartii* (Sacc.) Petch, commonly used for the management of root grubs, cutworms and termites, were tested for their pathogenicity against eggs of *M. incognita* in the laboratory.

The fungi were tested against fresh eggs obtained[3] from egg masses collected from infested roots of tomato following petri dish bioassay technique[4]. Spores of respective fungi were sprayed on 2% water agar plates at the rate of 10^7 conidia/plate; each plate was previously charged with 25 eggs of the nematodes. In control plates, only sterile distilled water was sprayed. Treatments were replicated four times using completely randomized design. Plates were incubated at 27 ± 1°C and fungal

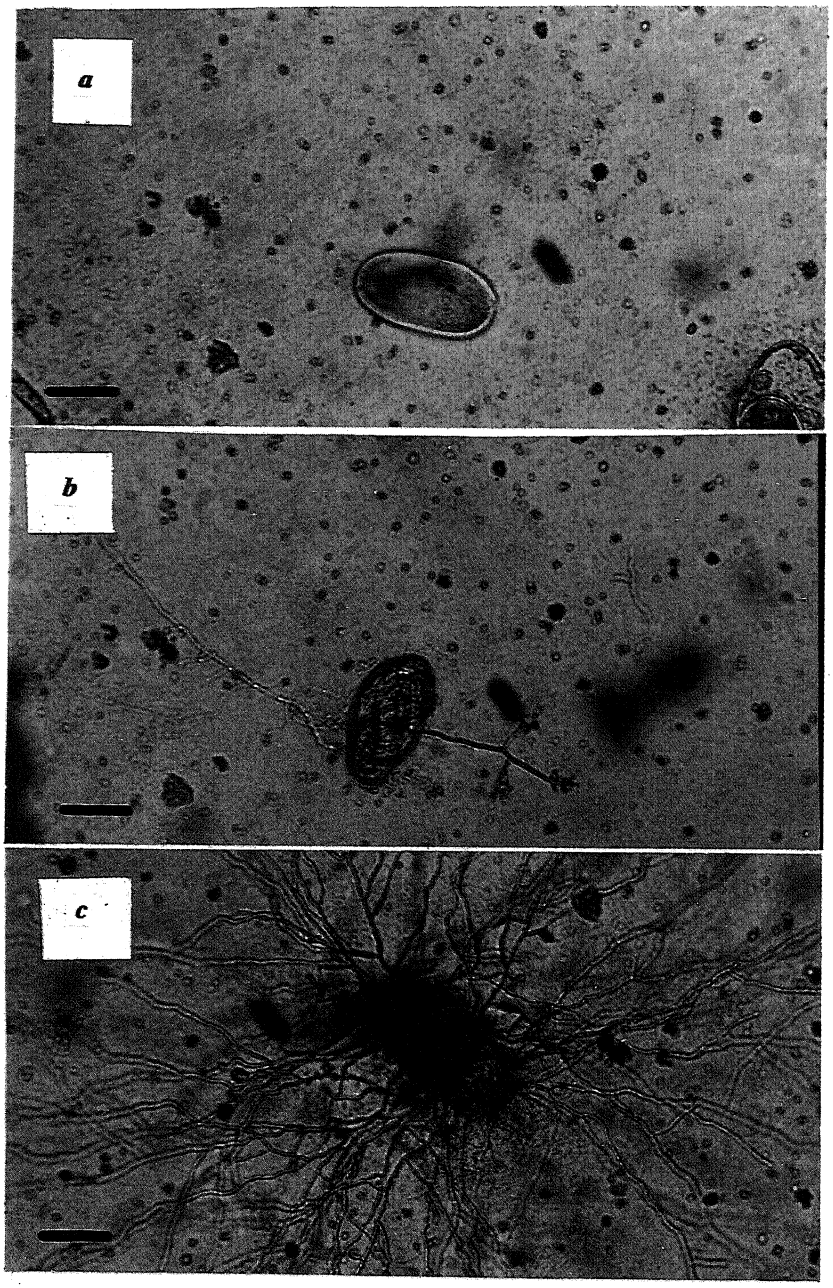


Figure 1 *a–c.* Parasitism of *M. incognita* eggs by entomopathogenic muscardine fungi (bar = 25 µm). *a*, Healthy egg of *M. incognita*; *b.* Egg parasitized by *Beauveria brongniartii*; *c.* Egg parasitized by *Metarhizium anisopliae.*

proliferation and colonization of nematodes eggs were recorded at 48 h intervals for two weeks. Hatching of larvae from eggs was also recorded from treated as well as control plates. Fungus-induced mortality in the eggs was confirmed under research microscope.

Results of plate assay clearly showed that both muscardine fungi have parasitized *M. incognita* eggs (Figure 1). *M. anisopliae* caused mortality in eggs ranging from 28.0 to 32.0%, with an average of $29.3 \pm 2.3\%$, while *B. brongniartii* caused 28.0 to 48.0% mortality in eggs, with an average of $38.7 \pm 2.0\%$. The average egg hatching in *M. anisopliae* and *B. brongniartii* treated sets was 56.0 ± 4.0 and $42.6 \pm 4.6\%$, respectively. A negative impact on egg hatching was observed due to fungus parasitism in treated sets compared to untreated control sets, where average $88.0 \pm 4.0\%$ normal hatching was recorded. This is the first report of *M. anisopliae* and *B. brongniartii* pathogenicity against root-knot nematode eggs in India. However, very recently, the pathogenicity of *Beauveria bassiana* has been reported against *M. incognita* from India[5].

Therefore, this study indicates that muscardine fungi, which are beneficial for management of soil insect pests, can also simultaneously be used for management of root-knot nematodes attacking agricultural crops. A complex of pest attack is very common in sandy loam type of soils, e.g. attack of white grub and *M. javanica* on groundnut and cutworms and *Meloidogyne* spp. on potato and tobacco in Gujarat. Further studies with these fungi on complex of nematode and insect under field conditions are warranted.

1. Bhatti, D. S. and Singh, Inderjit, in *Nematode Pests of Crops* (eds Bhatti, D. S. and Walia, R. K.), CBS Publishers, Delhi, 1992, pp. 5–13.
2. Stirling, G. R., *Biological Control of Plant Parasitic Nematodes*, CAB International, Wallingford, 1991, pp. 282.
3. Hussey, R. S. and Barker, K. R., *Plant Dis. Rep.*, 1973, **57**, 1025–1028.
4. McInnis, T. M. and Jaffe, B. A., *J. Nematol.*, 1989, **21**, 229–34.
5. Mishra, C., in Proceedings of the National Symposium on An Appraisal of the Nematode Management with Ecofriendly Approaches and Biocomponents, IARI, New Delhi, 24–26, March 1995, pp. 104.

R. V. VYAS
H. R. PATEL
B. A. PATEL
D. J. PATEL

Department of Nematology,. Gujarat Agricultural University, Anand 388 110, India

Intellectual property rights and biological resources: Specifying geographical origins and prior knowledge of uses

Madhav Gadgil and Preston Devasia

Intellectual property rights

Humans put to use a tremendously wide variety of other living organisms. Given the modern understanding and capabilities of manipulation of genetic material, this range now potentially encompasses all of living organisms. This has prompted an interest in claiming intellectual property rights (IPRs) over knowledge pertaining to uses of living organisms—not only novel organisms created through genetic manipulation, but even those merely isolated from nature. IPRs are also claimed over a variety of products of living organisms modified through human agency to varying degrees[1].

Living organisms have of course been genetically modified for millennia by farmers selecting seeds. A variety of biological products have also been used in applications ranging over cosmetics and pesticides to drugs. While such usage and knowledge may have been kept limited amongst certain clans or communities in the past, formal claims of IPRs are a recent development. In 1873 Louis Pasteur was awarded a patent on a yeast culture. Subsequently, many patents have been awarded on 'non-naturally occurring compositions of matter' based on living organisms designed for a variety of uses such as drugs and pesticides. Seeds began to receive patent-like protection in 1930s in the US and a genetically modified organism, a synthetic, plasmid injected *Pseudomonas* capable of degrading crude oil was awarded a US patent in 1980. In agreeing to the award of this patent, the US Supreme Court remarked that the US congress intended the statutory subject matter of the Patent Act to 'include anything under the sun that is made by man'. By now IPRs in the US extend to plasmids, vectors, cell lines, microorganisms isolated from nature, strains of cultivated plants and man-made animal forms such as polyploid oysters and transgenic mice[2]. Other countries are beginning to follow suit, and a greater and greater range of materials derived from living organisms, processes involving living organisms as well as self-reproducing living entities are now subject to IPRs.

The General Agreement on Trade and Tariff (GATT) aiming to reduce distortions and impediments to international trade includes an agreement on Trade-Related aspects of Intellectual Property Rights (TRIPS). TRIPS addresses itself to the issue of IPRs over living organisms and products and processes derived from them. It stipulates that members must accept patenting of microorganisms and microbiological processes, as well as protection of plant varieties either by patents or by an effective *sui generis* system or by any combination thereof. TRIPS presently permits members to exclude from protection plants (other than plant varieties) and animals and essentially biological processes for the production of plants and animals[3]. It is however likely that protection would be extended to these areas as well when GATT next comes up for review in 1999.

Convention on biological diversity

Concern for access to living resources, and their conservation has grown hand in hand with the rapidly expanding scope for use of living organisms towards human ends. These concerns led to the Convention on Biological Diversity (CBD) that came into force in December 1993, and has a current membership exceeding 125 countries. CBD aims to promote conservation of biological diversity, sustainable use of its components and equitable sharing of the resultant benefits. Till CBD living organisms were considered a common heritage of the humankind, this convention accepts them as a sovereign property of the nation states[4]. But living organisms are widely distributed over the world especially so since 1500 AD with rapid development of transoceanic transport. CBD therefore specifically recognizes 'countries of origin' of living resources and agrees to certain special treatment to such countries of origin, in particular those that belong to the categories of developing countries and small island states. All member countries of CBD agree to facilitate access to biological resources of which they are sovereign owners. In return the Convention promises developing countries of origin of such resources preference in location of research and development activities, favourable terms for transfer of technology and assistance towards development of technological capabilities. It promises such treatment not only on behalf of national governments, but private enterprises as well. It of course stipulates that this shall be on mutually agreed terms and in consonance with the IPR agreements (Articles 16 and 18, Convention on Biological Diversity[4]).

CBD not only recognizes special claims of developing countries of origin of living resources, but recognizes the role of indigenous communities also, in particular their womenfolk in conservation and sustainable utilization. It calls for a recognition of their conservation practices and their knowledge of manifold uses of living resources and stresses the need for equitable sharing of benefits of utilization of these resources with the indigenous communities (Article 8j, Convention on Biological Diversity[4]).

Two approaches

GATT (including TRIPS) and CBD thus represent two significant and separate approaches to the utilization of living resources of the earth to human ends. GATT aims to promote global economic growth by removing impediments to international trade. TRIPS aims to promote such economic growth through fostering technological innovation by guaranteeing on a uniform, international basis some degree of exclusivity to inventors on their inventions in order to be assured of a reasonable profit and to justify risks of development. CBD would also contribute to such global economic growth by

facilitating access to living resources. But CBD equally focuses on conservation and sustainable use, coupled to equitable sharing of benefits of such use. In its attempts to promote conservation, Article 22 of CBD stipulates that the obligations of the contracting parties under the convention are to take precedence over their obligations arising from other international agreements where the latter obligations would cause serious damage or a threat to biological diversity. In its attempts to create a stake for developing countries in conservation of biodiversity CBD stresses the need to share with them benefits of use of living resources, including the need to share development of and transfer of technology. These provisions have been criticized, in particular by US representatives as a serious interference with the intellectual property principles embodied in TRIPS and GATT. It was primarily on these grounds that the US has remained outside CBD, although almost all other major industrial and developing countries have joined it[5].

Development and conservation

In the emerging world order it is now necessary to flesh out the provisions of CBD to clarify their relationship with IPR-related regulations and to promote CBD's vital objective of long term conservation of biological diversity. The CBD provisions in this context attempt to create a stake in conservation of biodiversity for biodiversity-rich developing countries and indigenous communities of such countries living close to the natural world. CBD acknowledges that development to alleviate poverty is necessarily a high priority for these countries and communities. Biodiversity conservation would therefore involve serious opportunity costs for them, for which they must be compensated in order to motivate them to participate in conservation efforts. The precise mechanisms to achieve this must now be worked out.

Need for recognition

Such an effort would require as a first step appropriate recognition of countries of origin of specific elements of biodiversity, and of pertinent knowledge and conservation efforts of indigenous communities. While the details of transfer of financial resources and technology that such an effort would entail are likely to be controversial, the desirability of according due recognition to the countries of origin and indigenous communities should be a universally acceptable proposition.

As of today, there are no formal mechanisms to accord such recognition to countries of origin of genetic resources, or to the indigenous communities for their knowledge and practices of sustainable use. CBD visualizes all member countries organizing an effort at inventorying and monitoring biodiversity and developing appropriate information systems. This is however a massive undertaking and would take considerable time, more so because many diversity-rich developing countries have inadequate scientific and technical capabilities to undertake such an exercise. Such an exercise may also leave out of its ambit documentation of knowledge and practices of sustainable use of various elements of biodiversity by indigenous communities. It is therefore necessary to work out additional mechanisms of awarding such recognition to countries of origin and to indigenous communities.

Patent specifications

We suggest that appropriate augmentation of the requirements for the specification accompanying applications for patents and other IPR protection would be one such very useful mechanism. The four universally accepted requirements for the patentability of any invention include: novelty, nonobviousness, usefulness and adequate specification. Every patent application must disclose details of the invention with the help of a patent specification that describes an invention in sufficient detail to: (a) allow the patent office, and other interested parties to assess its claims of novelty, nonobviousness and usefulness; and (b) allow a person skilled in the art to reproduce it. Such a specification or disclosure statement has to accompany, or shortly follow a patent application[2,6]. Patent legislation has historically developed in relation to mechanical and chemical innovations. The specification requirements for innovations involving the use of biological materials are therefore only now beginning to evolve, primarily through judicial interpretations. Such interpretations have resulted in the requirement of deposition of the appropriate living material in an internationally recognized repository such as the Fermentation Research Institute of Japan or the European Collection of Animal Cell Cultures in the United Kingdom[2]. But there is no clearcut understanding today that the specification must include fuller details regarding the biological material such as the country of origin. Nor is there so far any formal acknowledgement of the need to accord recognition to community-based knowledge. However, many patent specifications do voluntarily provide such information. Thus European Patent 0010061 A1 800416 concerning novel pharmaceutical preparations containing an extract of the mollusc *Perna canaliculus* states it to be found on the shores of New Zealand[7]. In the case of American Patent 05298251 (1992) regarding fungicides derived from neem oil there is a mention of prior knowledge of communities regarding its uses[8].

Evidently, a further, quite marginal increase in the formal requirements of the specification would admirably serve the purpose of according recognition to the countries of origin and indigenous communities involved in sustainable use of the concerned biological resources. The proposed requirement should not only cover inventions which result in living material, such as a microorganism isolated in a culture, seed of a variety of cultivated plant, or a transgenic animal. It should also cover other products dependent on biological sources, such as alkaloids derived from plants and used as therapeutic agents or pesticides. Furthermore, the requirement should also extend to products that are essentially derived from biological sources, for instance, a synthetic molecule that differs a little from but has the same aroma as sandal (*Santalum album*) and was inspired by it. In all these cases, the specification should include a clear mention of the biological source material, the known country or countries of origin of that material, and all known information pertaining to knowledge and practices of sustainable use of that biological source material by indigenous communities in the country or countries of origin. In this context, it is essential to clarify the meaning of the term country or countries of origin. We suggest that the country/countries of origin be defined as those countries in which the biological material is known to have occurred as components

of natural biological communities, or under cultivation/domestication prior to the modern age of global transport beginning 1500 AD.

Geographical distribution

Ascription of a country or countries of origin to any genetic resource is evidently related to the patterns of geographical distribution of living organisms. Certain groups of organisms are apparently ubiquitous in their distribution. This is likely to be the case with spore-forming microorganisms. Thus most species of the slime mould genus *Dictyostelium* range over all continents and from tropics to arctics[9]. However, such wide ranging species are likely to harbour considerable intraspecific genetic variation within their populations. It is therefore possible that specific genetic variants may be confined to one or few countries of origin. Other microorganisms commensalic with higher plants or animals may have more restricted distributions. Thus the ectomycorrhiza *Tricholoma marsutaki* is known only from the host *Pinus densiflora* in certain specific habitats in Japan[10]. A few higher organisms are ubiquitous, but most have distributions limited to one or few countries or had such distributions prior to the suggested cut off date of 1500 AD. Many species of even highly mobile organisms like birds are endemic to just one country. Thus of 969 species of Indian birds, 69 are endemic, while of a total of 1519 species of Indonesian birds as many as 258 occur in no other country. Levels of endemism are higher in other, less mobile groups of organisms. Thus of 206 Indian species of amphibians as many as 110 are endemic, and of 270 Indonesian species, fully 100 are endemic[11].

Multiple lineages

The question of country of origin would also be complicated in case of organisms produced through human intervention from multiple lineages. Thus the high-yielding rice variety IR 64 produced at the International Rice Research Institute in Philippines is based on genes from 20 landraces coming from eight different countries namely China, India, Indonesia, Korea, Philippines, Thailand, USA and Vietnam[12]. The situation is apt to become even more complex in days to come. Already there exist transgenic organisms like *Nicotiana tabacum* with an endotoxin-producing gene from *Bacillus thuringiensis*[13]. *N. tabacum* may be assigned to several tropical American countries as countries of origin, but *B. thuringiensis* has a worldwide distribution. How to deal with issues of geographical origin of such transgenics clearly calls for further thinking.

Public scrutiny

Specifications accompanying patent or other IPR applications are open to public to varying degrees and after varying intervals following the filing of an application. It is suggested that this part of the specification relating to the biological source, country/countries of origin and knowledge and practices of sustainable use by indigenous communities be open to full pubic scrutiny immediately after the filing of the IPR application. This would permit countries with possible claims as countries of origin, and as repositories of public knowledge and practices of sustainable use to examine the patent applications and make any submissions that they may wish to make. It would also permit other parties, such as NGOs interested in furthering the interests of indigenous communities to make such submissions.

Inclusion in CBD and TRIPS

Inclusion of such specification in the patent application of course does not commit the inventor or others working the innovation to share the benefits arising therefrom in any way. That would obviously remain a matter of further negotiation. However, the specification would immediately help motivate potential country/countries of origin to document their own biological resources and to participate in *in situ* conservation efforts. It would also promote efforts to document and help perpetrate knowledge and practices of sustainable use by indigenous communities. Both these are important objectives of the CBD, with very broad-based acceptance at the international level. We therefore suggest that such specification should be accepted not only as an elaboration of CBD by all countries party to it, but also be made a part of the TRIPS under the GATT.

1. Gollin, M. A., in *Biodiversity Prospecting: Using Genetic Resources for Sustainable Development* (eds Reid, W. V., Laird, S. A., Meyer, A. A., Gamez, R., Sittenfeld, A., Janzen, D. H., Gollin, M. A. and Juma, C.), World Resources Institute, Washington, D.C., 1993, pp. 159–197.
2. Jong, S. C. and Birmingham, J. M., *Adv. Appl. Microbiol.*, 1990, **35**, 255–293.
3. Article 27 of the Trade-Related Aspects of Intellectual Property Rights, General Agreements on Tariffs and Trade – Final Text of Uruguay Round, 1994, World Trade Centre, Bombay.
4. Convention on Biological Diversity, 1992, United Nations Environment Programme, Nairobi.
5. Morrow, J. D. and Britt, K. R., *Intellectual Property and Biodiversity*, 1992 (personal communication).
6. Morrow, J. D., Biotechnology Patent Conference Workbook, American Type Culture Collection, Maryland, 1993, pp. 73–80.
7. McFarlane, S. J. and Croft, J. E., European Patent 0010061 A1 800416.
8. Locke, J. C., Walter, J. F. and Larew, III, H. G., US Patent 05298251, 1992.
9. Cavender, J. C., *Am. J. Bot.*, 1976, **63**, 71–73.
10. Harley, J. L. and Smith, S. E., *Mycorrhizal Symbiosis*, Academic Press, New York, 1983.
11. *Global Biodiversity*, compiled by World Conservation Monitoring Centre (ed. Groombridge, B.), Chapman and Hall, London, 1992.
12. Swaminathan, M. S., in *Farmers' Rights and Plant Genetic Resources—Recognition and Reward: A Dialogue* (ed. Swaminathan, M. S.), Macmillan, Madras, 1995, pp. 200–241.
13. Barton, K. A., Whiteley, H. R. and Yang, N. S., *Plant Physiol.*, 1987, **85**, 1103–1109.

ACKNOWLEDGEMENTS. We thank N. S. Gopalakrishnan, K. Gopinath, D. J. Bagyaraj, V. Nanjundiah, A. Seetharam and several participants at the first meeting of the Subsidiary Body on Scientific, Technical and Technological Advice to the Convention on Biological Diversity for most useful discussions.

Madhav Gadgil is in the Centre for Ecological Sciences, Indian Institute of Science, Bangalore 560 012, India and Biodiversity Unit, Jawaharlal Nehru Centre for Advanced Scientific Research, Jakkur P.O., Jakkur, Bangalore 560 064, India; and Preston Devasia is in Biodiversity Unit, Jawaharlal Nehru Centre for Advanced Scientific Research, Jakkur P.O., Jakkur, Bangalore 560 064, India.

Developments and trends in inductively coupled plasma mass spectrometry and its influence on the recent advances in trace element analysis

V. Balaram

A brief introduction to the various instrumental methods such as atomic absorption spectrometry, X-ray fluorescence spectrometry, neutron activation analysis, inductively coupled plasma atomic emission spectrometry, thermal ionization mass spectrometry, etc. are presented highlighting their relative merits and demerits. The history and developments of inductively coupled plasma mass spectrometry (ICP–MS) and its advantages and limitations over other multi-element instrumental techniques are reviewed. Extended capabilities by hyphenating ICP–MS to various other well-known sample introduction techniques such as flow-injection, electrothermal vapourization, chromatographic methods and laser ablation are discussed in brief. The recent development of high resolution multi-collector double-focusing magnetic mass spectrometer with inductively coupled plasma at atmospheric pressure as source is also discussed. Some of the areas where more developments can be expected in future are suggested.

RECENT developments in basic material research, environmental studies, quality control protocols in most industries, etc., have placed new demands on the field of metal analysis. Detection and accurate estimation of almost all the elements in the periodic table at ng/ml or μg/ml range has become an important requirement in the fields of both basic and applied research[1–4]. For example, in the environmental analysis field, the number of regulated elements is becoming more and more and the required levels are moving down to lower levels. Regulations that are being established by many countries for the inorganic contaminants in water and food-related materials, the increased awareness of the importance of trace elements in foods and the element speciation have created demands mainly for increased sample throughput and for lower levels of detection. Earth scientists working on geochemical problems need to determine relative and absolute abundance of the elements present even at trace and ultra-trace levels in a variety of rocks and minerals, with the object of discovering principles governing their distribution and migration[5]. Precise estimation of isotope ratios of certain elements such as Sm, Nd, Pb, U, etc., at very low concentrations is important for applications such as radiogenic dating of rocks, ore deposits and in very accurate estimation of certain elements by isotope dilution method. There is also a need to characterize a variety of materials for trace metal impurities in several industries. The semiconductor industry, for example, requires determination of several trace elements in pure materials. As the density of the circuitry increases and the geometry shrinks, the need for pure and purer materials increases. The R&D laboratories at the chip manufacturers are working on the needs of one or two generations into the future because the pace of development accelerates so fast. Most of the other industries need their finished products to be analysed for the presence of trace metal impurities. For example, analysis of trace elements in copper is of great significance in electronic industry, where properties of copper may be severely affected by the metal contaminants at levels even below part per million concentrations (μg/g). The need for purer and purer reagents for the R&D laboratories, very often requires checking of the common mineral acids and other laboratory reagents for the contamination by certain elements at ultra-trace levels (for example, Rb, Sr, Sm, Nd, Pb, U, etc., for geochronological studies). Finally, the role of trace elements in human health and disease[6] is an area of immense interest for it is known that over a quarter of elements in the periodic table are essential to life. Thus, highly precise and accurate estimation of several elements at very low concentrations is required in a variety of fields such as geological, industrial, nuclear, environmental and biomedical. Most often, the sample volumes are relatively small and analytical requirements relatively high requiring very highly sensitive instrumental techniques for their analysis.

V. Balaram is in the National Geophysical Research Institute, Hyderabad 500 007, India.

Methods for the analysis of elements

From the turn of this century there has been an analytical revolution with succession of a series of analytical

instruments with multielement capabilities. The dominant instrumental analytical techniques available currently for qualitative identification and quantitative estimation of most elements in the periodic table include atomic absorption spectrometry (both flame and graphic furnace AAS), X-ray fluorescence spectrometry (XRF), neutron activation analysis (both instrumental and radioanalytical NAA), thermal ionization mass spectrometry (TIMS), spark source mass spectrometry (SSMS) and inductively coupled plasma atomic emission spectrometry (ICP–AES). Most of these are capable of carrying out rapid multielement analysis at μg/ml levels, except for AAS which can detect and determine one element at a time. Other disadvantages of AAS (both flame and flameless) include the limited concentration range (requiring multiple dilutions to get a signal of appropriate intensity on scale) and the sequential nature of most systems. The fact that only one or two elements can be determined at a time is the major reason which caused its decline in favour of ICP–AES, which was made available in the mid-1970s. Though ICP–AES provides multielement detection for most elements at sub μg/ml level, there are numerous limitations resulting from complex spectra and matrix interferences. XRF which started in the mid-1950s is a popular analytical technique in both research and industry for the estimation of major and minor elements in several matrices such as rocks, alloys, etc. This technique suffers from drawbacks such as, interferences, very high detection limits for several trace elements in the above matrices, and also this technique cannot handle solutions. One of the main strands of analytical advancement was the development of neutron activation analysis (NAA). Although the principles of NAA had been understood since the publication of Hevesy and Levi's[7] work in 1936, the full potential of the technique was realized only during the 1960s after the development of high resolution gamma-ray detectors[8]. It is a nondestructive technique and offers multielement capability, together with relatively good levels of sensitivity (μg/ml levels). The need for a nearby reactor often precludes the consideration of NAA. If such a facility exists, the technique can be valuable. However, a greater disadvantage of NAA is the production of radioactive waste apart from having matrix interference effects. However, the radiochemical separation (RNAA) overcomes many matrix interference problems and offers low detection limits (sub ng/ml levels). But it involves many long drawn separation procedures which may carry the potential of delivering a high dose of radiation to the analyst. Thermal ionization mass spectrometry (TIMS) is a well-established method for measuring both radiogenic and nonradiogenic isotope ratios. Isotope ratio measurement of elements such as neodymium and strontium is routinely accomplished to high levels of precision and accuracy. But certain elements cannot be analysed by this technique as the efficiency of thermal ionization decreases with increase in their ionization potentials. Further disadvantages of thermal ionization include, the extensive sample preparation and prolonged analysis times. TIMS is also applicable to elemental analysis. Very high precisions are attainable using the isotope dilution method, though this is limited to elements with more than one naturally occurring isotope. Though in recent years multicollector systems for parallel collection of several isotopes have been perfected, and offer enhanced efficiency and drift compensation, TIMS is not suited to the routine analysis of samples in large numbers.

During the last 15 years, considerable attention has been paid to the development of yet another and more powerful analytical tool called inductively coupled plasma mass spectrometry (ICP–MS) by the combination of plasma ion source with mass spectrometry techniques, which has revolutionized elemental analysis and has become widely accepted as a powerful tool for trace and ultratrace elemental and isotopic analysis. Thus, evolution of instrumentation continues with the latest method being ICP–MS.

History and development of ICP–MS

Gray[9] was the first to demonstrate the potential of the plasma source mass spectrometry in 1975. He used a direct coupled plasma (DCP) and showed that strong signals could be obtained with low-background levels, so that detection of elements in μg/ml region was possible in aqueous solutions. Although these results were encouraging initially, they also revealed the limitations of the DC plasma used. The gas temperature of about 3500 K was too low to provide adequate sample volatilization, dissociation and ionization of the elements, especially those with ionization energies above 8 eV. Microwave-induced plasma (MIP) was also rejected on the basis of its low-gas temperature and ICP was considered the most suitable. These historic observations by Gray followed by the pioneering work[10–14] by him at the University of Surrey along with Alan Date, at British Geological Survey, UK; R. S. Houk, from Iowa State University at Ames, USA and Don Douglas, at Sciex, Canada in the subsequent years have brought about the development of this powerful analytical instrument. The major innovation in the design of ICP–MS is the interface between the ICP and the quadrupole mass spectrometer. Here, ions produced in the plasma at atmospheric pressure are transported into the quadrupole mass analyser held under ultra-high vacuum through a differentially pumped vacuum system which is not technically simple. The first commercial instruments were introduced in 1983 by two firms, namely V.G. Isotopes (now FI Elemental Analysis, UK) and Sciex Inc. (now Sciex–Perkin Elmer, USA). Perhaps a reflection of the success

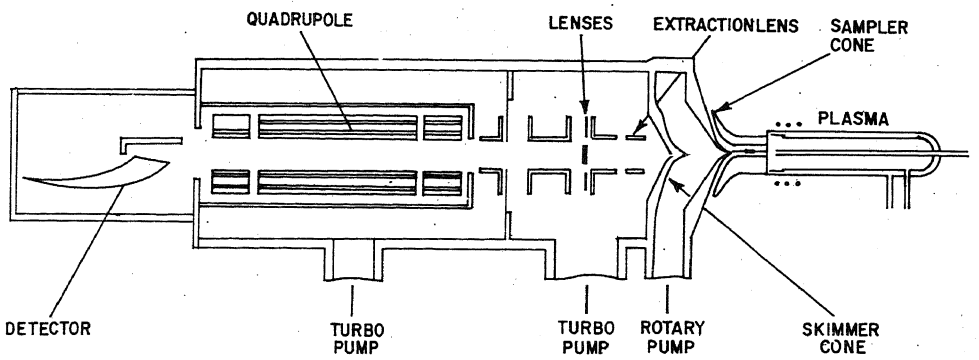


Figure 1. A schematic of plasmaquad system.

is that more than ten firms are manufacturing and marketing this particular instrument world-wide today. Efforts are also on at BARC[15], Bombay in India to produce this instrument indigenously. Though there are several hundreds of instruments installed in chemical laboratories all over the world, there are only two ICP–MS instruments in India at present. One instrument (Elan 250, Sciex, Canada) is at the Indira Gandhi Centre for Atomic Research, Kalpakkam and the other (VG Plasmaquad, FI Elemental, UK) at the National Geophysical Research Institute, Hyderabad. Though these two instruments were procured mainly for the analysis of nuclear and geological samples respectively, they have also been continuously utilized by several other groups in the country working on a range of disciplines such as geochemical, environmental, water, biological and metallurgical fields since the last eight years.

Over the last decade, ICP–MS has been well established as a powerful tool for simultaneous multielement analysis. The low detection limits, high sample throughput, small sample quantities required for analysis, element versatility (major, minor, trace and ultra-trace) and isotopic capabilities are some of the important features which have made ICP–MS an excellent spectroscopic technique for the analysis of a variety of materials.

System outline and the principle

Configuration[16] of the instrument consists of three basic units: i) a conventional argon ICP operating at temperatures of 6000–9000 K with a nebulizer, spray-chamber, work coil and associated power supplies, ii) a conventional quadrupole mass spectrometer and associated data collection electronics, which permit rapid scanning of the selected mass range between 0 and 300 amu (atomic mass units), and iii) an interface unit consisting of two water-cooled nickel cones, each containing a small orifice at the centre, which allow sampling of the plasma gases and transfer of the ion-beam into the mass spectrometer.

The principle of ICP–MS can be understood with reference to the schematic diagram given in Figure 1. Samples in the form of solution are introduced through the peristaltic pump at a rate of about 1 ml/min into the central region of the plasma at atmospheric pressure with the help of a nebulizer and a water-cooled spray-chamber system. The sample is heated to a temperature of approximately 9000 K in the plasma, resulting in a series of processes involving desolvation, vaporization, dissociation, atomization and ionization[17] (Figure 2) in the 'analytical zone' of the ICP. At this temperature, chemical interference effects are insignificant, and essentially all atoms are ionized to a very large extent (majority of the elements in the periodic table are expected to ionize to the extent of about 90% in ICP). The efficiency of ionization in the argon plasma as calculated by the Saha equation[18] shows that the formation of singly charged ions is very efficient in argon plasma and over 54 elements are expected to be ionized with an

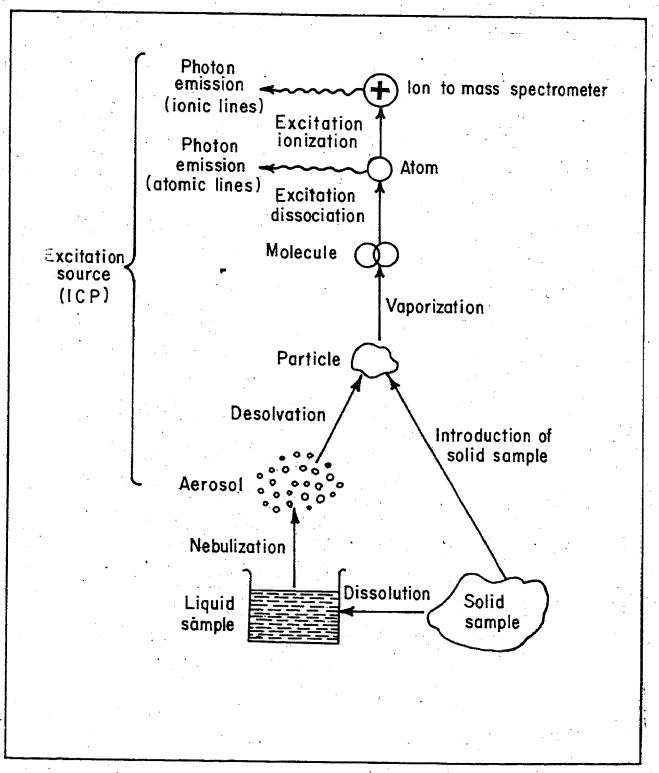


Figure 2. The chain of events leading from sample to the formation of ions.

efficiency of 90% or more[16]. A fraction of the positively charged ions produced in the plasma is transported through a narrow aperture of the sampler cone[19]. Here the plasma beam expands at supersonic speeds because of vacuum conditions (1 mbar) and is sampled by the skimmer cone. The ions and the plasma gas pass into a further high vacuum region (1×10^{-4} mbar) and pass through a system of electrostatic lenses which extract the positively charged ions and focus them in a form suitable for transmission through the quadrupole mass spectrometer which is kept in a very high vacuum region (2×10^{-6} mbar). The mass spectrometer transmits only ions of a particular mass to charge ratio, any neutral species accompanying the plasma gases would be taken away from the system by the vacuum pumps. This minimizes the ion–molecular collisions in the system. After mass isolation, the filtered ions pass into the detector system. An ion-detector system consisting of a channel electron multiplier is used in pulse-counting mode, to register the transmitted ions and generate a signal for every ion impinging upon its surface which will be amplified further. Since the number of registered ions from a given isotope, of course depends directly on the concentration of the relevant isotope (element) in the sample, quantification is straightforward. Design of the ion-optics makes the sampled ions to take a curved trajectory around a central photon stop eventually to be counted by an off-axis detector. This arrangement supresses photon noise as the detector is also sensitive to photon radiation from ICP[16] and produces a low background signal of only around 25 counts per second as against a count rate of about 1 lakh for a 50 ng/ml solution of a monoisotopic element like rhodium. Detection limits, mainly influenced by the relative abundance of the isotope selected for analysis for most elements are currently in the range 0.01 to 0.1 ng/ml (for example, the detection limits of rare earth elements in Table 4). These detection limits are obtainable under multielement operating conditions. For many elements, even better detection limits can be obtained if operating conditions for the plasma, ion optics and mass analyser are optimized exclusively for one element or a few elements of similar mass.

Performance characteristics and applications

The instruments constructed to date employ quardrupole mass analysers which yield essentially better than unit mass resolution over a mass range up to 300 amu. The analyser can scan very rapidly the entire useful mass range (0–240 amu) or a selected mass window of interest. A mass spectrum is normally obtained by many such repeat scans (say 100) which may take around 60 sec. This process of mass scanning is very rapid, and the instrument can obtain a spectrum for the entire mass

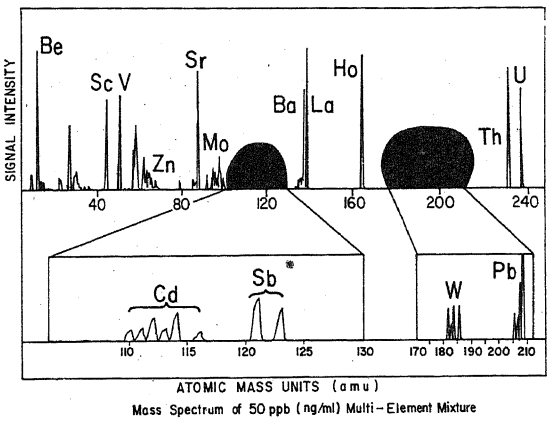


Figure 3. Mass spectrum of 50 ppb (ng/ml) multielement mixture.

range of ^{7}Li to ^{238}U in about 1 min. One such spectrum is shown in Figure 3 for a 50 ng/ml multielement solution, the position of the peak in the spectrum reveals the nature of the element (isotope) in the sample and the height of the peak its concentration. By comparing the mass peaks with those observed for a standard solution run under identical conditions, exact concentrations of different elements in any sample can be precisely calculated after making proper blank and interference corrections by an on-line computer. In place of mass scanning, peak jumping can also be used for data acquisition. Despite early claims that ICP–MS was relatively free from interferences[12], many workers noticed several kinds of interferences such as matrix effects either exclusively or in conjunction with a particular application[20, 21]. Evans and Giglio[22] have presented a comprehensive review of interferences in ICP–MS. Recently, Burton and Horlick[23] developed a graphically oriented data base of spectral interferences in ICP–MS, such as isobarics, oxides, doubly charged species, background- and matrix-dependent species, etc. This program is called 'MS Interview' which runs on a Macintosh Computer and helps the user to broadly assess the interference problems in ICP–MS. The most important limitation of this technique is nontolerance to high total dissolved solids in the sample solution. More than 0.4% of total dissolved solids cause clogging of the sampler and skimmer cones. For all practical purposes running of 0.1% solutions is recommended for obtaining reasonable long-term stability[24]. Others include the potential interferences from the background and sample matrix, periodic drift in the signal, space-charge and mass discrimination effects, etc. A few of these problems are minimized by the use of internal standardization[25], matrix-matching standards[26], addition of chemical modifier[27], isotope dilution[28], etc. during analysis.

The advantages of ICP–MS include: (i) very low detection limits which are considerably better than XRF and ICP-AES[26], (ii) much wider elemental coverage than by AAS or INAA, (iii) capability of doing multielement

Table 1. Trace element data on hornblende mineral reference sample, JH-1 (Geological Survey of Japan) obtained by ICP–MS (at NGRI) in comparison with the values obtained by other popular analytical techniques together with certified values

Element	Values (μg/g)	Technique	Certified values*	Element	Values (μg/g)	Technique	Certified values*	Element	Values (μg/g)	Technique	Certified values*
Sc	82	XRF		Y	12	ICP–AES	—	Tb	0.52	INAA	—
	73	ICP–AES			13	FE			0.46	ICP–MS	
	74	FE			14	XRF					
	81.4	INNA			12.22	ICP–MS					
	71.27	ICP–MS									
V	225	FE	231	Zr	42	FE	—	Dy	2.6	GRAV	—
	242	XRF			51	XRF			2.89	ICP–MS	
	272	ICP–AES			86	ICP–AES					
	229.93	ICP–MS			51	ICP–MS					
Cr	617	ICP–AES	630	Nb	3.6	FE	—	Er	1.2	GRAV	—
	625	GRAV			4.6	XRF			1.2	ICP–MS	
	630	INAA			24	ICP–AES					
	669	XRF			3.36	ICP–MS					
	623.71	ICP–MS									
Co	51	GRAV	53	Cs	1.14	INAA	0.5	Yb	1.15	FE	—
	51.6	INAA			0.82	ICP–MS			1.23	INAA	
	55	XRF							1.40	ICP–MS	
	52.11	ICP–MS									
Ni	55	FE	56	Ba	106	GRAV	—	Lu	0.20	INAA	—
	56	XRF			109	XRF			0.13	ICP–MS	
	93	ICP–AES			169	ICP–AES					
	56.91	ICP–MS			112.37	ICP–MS					
Cu	9	GRAV	9	La	7.5	ISE	—	Hf	1.46	INAA	—
	11	XRF			7.6	INAA			1.16	ICP–MS	
	18	ICP–AES			8.18	ICP–MS					
	9.66	ICP–MS									
Zn	67	XRF	62	Ce	17	GRAV	—	Ta	0.24	INAA	—
	92	INAA			17.20	INAA			0.22	ICP–MS	
	130	ICP–AES			19.23	ICP–MS					
	64.26	ICP–MS									
Ga	9	XRF	—	Nd	11.4	INAA	—	Pb	3	XRF	—
	7.36	ICP–MS			12	FE			2.56	ICP–MS	
					11.27	ICP–MS					
											—
Rb	12	ICP–AES	12	Sm	3	INAA		Th	1.54	INAA	
	14	FE			3.42	ICP–MS			1.57	ICP–MS	
	15.4	XRF					—				
	14.09	ICP–MS									—
Sr	152	FE	155	Eu	0.85	GRAV		U	0.6	INAA	
	154	XRF			0.90	INAA			0.46	ICP–MS	
	306	ICP–AES			0.94	ICP–MS	—				
	156.05	ICP–MS									

GRAV = Gravimetry; XRF = X-ray fluorescence spectrometry; ICP–AES = inductively coupled plasma atomic emission spectrometry; ICP–MS = Inductively coupled plasma mass spectrometry; FE = Flame emission spectrometry; INAA = Instrumental neutron activation analysis. *The values obtained by ICP–MS (at NGRI) are average of six values. Analytical data on ICP–AES, XRF, FE, INAA and GRAV are by N. Imai, Geological Survey of Japan (personal communication, 1994).

analysis, unlike AAS with which we can estimate only one element at a time[29], (iv) considerably faster sample throughput at a rate of one sample in 3 min for the analysis of over 30 elements[30] and (v) the unique capability of providing a very rapid semiquantitative survey analysis for over 70 elements in an unknown sample[31]. Fulford and Quan[32] have studied the basic characteristics of negative ions such as F^-, C^-, Br^-, I^- and S^- in ICP–MS in the negative ion mode and revealed that they do not originate directly from the plasma but rather from electron capture or reaction downstream of the free jet expansion. They also indicated the possibility of sensitive detection of halogens, especially fluoride, which makes this mode particularly attractive where halogens are of

Table 2. Concentrations of various trace elements (ng/ml) in NIST 1643c, a water reference sample, obtained by ICP–MS (at NGRI) in comparison with certified values

Element	ICP–MS value*(ng/ml)	Certified value† (ng/ml)
Be	22.91	23.2
B	118.75	119.00
V	32.07	31.4
Cr	20.29	19.00
Mn	35.79	35.1
Fe	105.34	106.9
Co	23.98	23.50
Ni	61.01	60.60
Cu	22.00	22.30
Zn	75.11	73.90
As	81.48	82.10
Se	12.49	12.70
Rb	11.30	11.40
Sr	266.26	263.60
Mo	102.54	104.30
Ag	2.13	2.21
Cd	12.18	12.2
Ba	50.02	49.60
Tl	8.02	7.90
Pb	34.70	35.30
Bi	12.40	12.00

*Average of three estimations.
†National Institute of Standard and Technology Certificate, Standard Reference Material 1643c, Trace Elements in Water, 1991.

Table 3. Concentrations of gold and some platinum group elements determined by ICP–MS (at NGRI) in a platinum ore reference sample, SARM-7 (Council for Mineral Technology, South Africa) in comparison with certified values

Element	Value obtained by ICP–MS (μg/g)*	Certified values† (μg/g)
Ru	0.41	0.43
Pd	1.55	1.53
Ir	0.06	0.074
Pt	3.67	3.74
Au	0.32	0.31

*Average of three determinations
†Govindaraju, K., *Geostand. Newslett.*, 1994, **18**, 1–158.

specific interest. Other advantages include isotopic analysis capability, relatively few interference effects, linear dynamic range for about 5–6 orders of magnitude, etc.

ICP–MS technique has been applied virtually to all types of samples, including geological[24,33], marine[34], biological[35], nuclear[36], freshwater[37], sea water[38], metallurgical[39], environmental[31], petroleum products[40], etc. There also have been a few instances of measuring trace elements by ICP–MS at pg/g levels in ultra-pure acids[41]. Although some samples like water could be aspirated directly, pretreatment is required in most cases. The trace element data obtained at our laboratory (Tables 1–3) on the hornblende mineral, water and precious-metal ore reference samples by ICP–MS demonstrate the potentials of ICP–MS in comparison with other popular analytical techniques. The methods used for sample preparation and other details are presented elsewhere[26,31,42,43]. The measurement precisions obtained on various trace elements are better than 5% relative standard deviation (RSD) with comparable accuracies.

Extending the capabilities of ICP–MS

ICP–MS was basically designed for multielement analysis of aqueous samples. In order to achieve accurate, reliable and sensitive results, especially when the concentrations of analyte elements in the original material or the prepared solution are too low to be determined directly by ICP–MS, or when matrix elements interfere with the determination, preconcentration and separation techniques are needed. In recent times, efforts have been successfully made to extend its capabilities by simple hyphenations to other techniques[44]. Some of them are discussed in the following:

Flow-injection method

Alternative sample introduction technique such as the flow injection analysis (FIA) when used in combination with ICP–MS is found to be very useful in agricultural, biochemical, clinical, pharmaceutical and environmental fields where smaller amounts of samples are available for analysis. This involves the injection of a discrete sample plug into a nonsegmented carrier stream. The sample moves into carrier stream from injection point to the ICP. A major limitation of ICP–MS is its sensitivity to the matrix (spectroscopic and nonspectroscopic interferences, clogging by solid deposition, etc.). When FIA preconcentration/separation techniques are employed, the analytes can be selectively separated from the bulk of the matrix[45] before injecting into ICP. The advantages include high throughput and minimal sample consumption. An example of effective utilization of FIA–ICP–MS is the estimation of very low levels of Cr, Ni and Pt in urine samples[46].

Electrothermal vaporization

Another alternative is to couple an electrothermal vaporization (ETV) unit to ICP–MS. ETV can be carried out using a metal filament or a heated graphite tube. Microlitre amounts of sample are injected into the tube, dried and then rapidly heated to approximately 3000°C to vaporize the sample, which is then transported to the ICP by using a carrier gas-flow. Rapid quadrupole scanning combined with multichannel data acquisition

allows the resultant transient signals to be captured and analysed[47]. This results in decrease of interference effects and improved sensitivity[48]. One interesting example of this technique is the determination of Os in geological materials[49] as it is very difficult to estimate Os in geological samples by solution nebulization.

Slurry nebulization

Although ICP–MS is designed for the analysis of solutions, it can be applied to solids if they are finely ground and suspended in water. Williams *et al.*[50] published a feasibility study of analysing slurry concentrations of 0.05 g per 100 ml of tetrasodium pyrophosphate as dispensing agent for finely ground (<3 μm) sample particles. Slurries of up to 1% w/v can indeed be nebulized using a high-solids nebulizer[51]. This approach is particularly useful for samples which are hard to dissolve[52].

Coupling with chromatographic techniques

Although many of the analytical techniques available today, including ICP–MS, are capable of determining elements at very low levels, most of them yield information only of total levels. Though ICP–MS can provide isotopic information, it cannot distinguish between different oxidation states of an element. At present there is an increasing demand to quantify the form or oxidation state of trace elements in a wide variety of samples. This is often termed metal speciation. Such information is very important to environmental scientists and toxicologists[6]. One of the most promising approaches to obtain trace metal speciation information is by combining an appropriate chromatographic technique to ICP–MS. Because of chromatography's separation power, coupling it to ICP–MS provides speciation capability. High-performance liquid chromatography (HPLC) is ideal as it employs solutions[44]. In general, the mobile phase must be selected carefully to prevent clogging of the interface, extinguishing of the plasma, etc. Low levels (a few per cent) of organic solvents in water can easily be tolerated by increasing the power and water-cooling the spray chamber (to condense and drain out as much solvent as possible). For higher levels, addition of oxygen to the plasma is required to burn out the carbon. By interfacing liquid chromatographic system with ICP–MS Huang and Jiang[53] have been able to accurately determine mercury species (total mercury and methyl mercury) in Dogfish Muscle reference material. Houk *et al.*[54] also used ion-pair reversed-phase HPLC to separate various cationic species of Hg and Pb in human urine.

A variety of high-efficiency nebulizers such as Mistral nebulizer[55], ultrasonic nebulizer[54], direct injection nebulizer[56], hydraulic nebulizer[57] and recycling nebulizer[58] have been successfully employed together with ICP–MS to realize the full potential of the technique. For example, memory effects associated with conventional pneumatic nebulization were found to be drastically reduced with direct injection nebulization[56].

Laser ablation

For the analysis of high-purity materials, such as high-pure metals, semiconductors, etc., sample treatment must be kept to an absolute minimum. This requirement has led to the development of a technique called laser ablation–ICP–MS (LA–ICP–MS). Gray[59] demonstrated the feasibility of laser ablation for solid sample introduction to ICP–MS for the first time in 1985. Laser ablation is an attractive alternative sample introduction technique to dissolution, particularly for handling geological materials[60]. A pulsed Nd:YAG laser is used to ablate solid samples held within an ablation cell. The laser is focused onto the surface of the solid sample, and controlled laser pulses of a millisecond duration vapourize the analyte material. The microparticulate cloud produced is carried by a stream of argon gas to the ICP torch and subsequently ionized within the plasma followed by mass spectrometric analysis. It is of particular importance to the nuclear, semiconductor, mineral and ceramic industries where information on several trace elements at μg/g and ng/g levels are needed without dissolving the sample. This technique can also provide very rapid semiquantitative analysis of materials of completely unknown composition[61]. LA–ICP–MS also offers the possibility of producing spatial information on element distribution at micrometer scale in solid materials[62]. This technique has also been applied to the characterization of spent nuclear fuels and to study the isotopic distributions of fission products and actinides in fuel pellets[63,64] and conductive (e.g. alloys powders) and nonconductive powder samples[65] (e.g. calcium sulphate and titanium dioxide powders for industrial use).

High resolution–ICP–multiple collector mass spectrometry (HR–ICP–MC–MS)

ICP–MS equipped with quadrupole mass analysers has been applied to the measurement of isotope ratios[60]. It is currently not possible to achieve precisions better than 0.1% RSD in routine analysis. This is not sufficient for geochronological studies and nuclear purposes. An unique instrument with ICP as source and incorporating a double focusing magnetic sector mass analyser equipped with multiple ion collectors has been recently developed[66]. The isotopic components were measured simultaneously which allowed high precision and high accuracy isotope ratio analysis (better than 0.01% RSD) which were comparable to those exhibited by TIMS[67]. It is also demonstrated recently that by combining laser ablation with HR–ICP–MC–MS, the isotopic composition

Table 4. Detection limits (3σ values) of rare earth elements obtained by ICP–MS, LA–ICP–MS and HR–ICP–MS in comparison with those obtained by other well-established multielement analytical techniques

Element	XRF (μg/g)*	ICP–AES (μg/ml)*	INAA (μg/g)*	ICP–MS (ng/ml)**	LA–ICP–MS (μg/g)†	HR–ICP–MS (pg/ml)‡
La	1.17	0.044	0.83	0.010	0.11	0.0014
Ce	1.10	0.145	2.50	0.015	0.13	0.002
Pr	1.03	0.065	—	0.010	0.13	0.001
Nd	0.87	0.085	7.70	0.050	0.12	0.0058
Sm	0.77	0.031	0.17	0.036	0.11	0.005
Eu	0.73	0.005	0.08	0.017	0.14	0.003
Gd	0.67	0.035	6.5	0.012	0.12	0.005
Tb	0.67	—	0.15	0.009	0.17	0.0008
Dy	0.63	0.011	—	0.024	0.14	0.002
Ho	0.60	0.010	2.83	0.010	0.16	0.0005
Er	0.57	0.019	—	0.028	0.14	0.002
Tm	0.57	—	0.57	0.009	0.17	0.0008
Yb	0.57	0.004	0.23	0.025	0.18	0.002
Lu	0.50	0.004	0.17	0.008	0.16	0.0009

*ref. 70; **ref. 24; †ref. 77; ‡ref. 74;–not available.

of solids can be measured directly to similar superior levels of precision and accuracy even for those elements such as W, Sn, Te, Mo, etc., which are difficult to ionize by thermal ionization[68]. High resolution ICP–MS offers detection limits better than 0.1 pg/ml in aqueous solutions for several elements. Table 4 gives the detection limits obtained by HR–ICP–MS for rare earth elements in comparison with those obtained by other multielement analytical techniques. We can see that the detection limits obtained by HR–ICP–MS are better by several orders of magnitude. This is a powerful analytical tool, especially when complex matrices like geological, ceramic, biological, sea water, etc. are investigated. Such matrices not only suffer from many spectral interferences such as isobaric overlaps, interferences from molecular ions, doubly charged ionic species, etc., which are well known but also from numerous unexpected spectral interferences which may additionally arise because of the complexity of these matrices and confuse the spectral interpretation. High resolution mass spectrometers keep the analyst free from most of the significant problems from spectral interferences in real samples while simultaneously improving sensitivity[69]. At least three firms, viz. FI Elemental Analysis, Winsford, UK (PlasmaTrace); Jeol Ltd, Tokyo, Japan, (JMS-Plamaxi1); and Finnigan MAT, Bremen, Germany (ELEMENT) have launched high resolution instruments during 1993.

Future directions

New developments in ICP–MS instrumentation and applications are advancing the understanding and the utility of the technique. The following are some of the areas where striking improvements can be expected in future.

i) During the nebulization of the samples, experience[70] has shown that approximately only one in one million ions sampled actually reaches the detector. This astonishing inefficiency of the present instruments suggests that further refinement of the design of the interface and ion-optics will improve the sensitivity further. Space-charge effects are thought to be a major cause of ion loss in ICP–MS. Briefly, as the beam leaves the skimmer cone it encounters the electric field of ion lens and the electrons are stripped away. The current from the ions remaining in the beam is far too high for the lens to tolerate so that the ion beam explodes, resulting in loss of ions. Active research is going on in this direction to minimize these effects. One firm (Hewlett-Packard, USA) has eliminated the need for a photon stop by mounting the quadrupole and detector off axis. An omega lens is used to bend the ion beam into the entrance aperture of the quadrupole. By this arrangement it is claimed that the new design has achieved a very high ion transmission with minimal space-charge effects, resulting in very low background and high sensitivity compared to the conventional systems[71].

ii) The normally operated nebulizers deliver only 1–2% of the sample to the plasma. This fraction can be increased roughly by an order of magnitude by using ultrasonic nebulizers as they produce smaller, more uniform droplets and is generally not prone to blockage and so is more efficient than pneumatic nebulizers[72]. Striking advances have already been made in this direction and some more may follow in future.

iii) Another effective way of minimizing the interference effects is to reduce the water load to the ICP by using the multiple desolvation[73] steps at –80°C for a continuous flow ultrasonic nebulizer. It has been observed that oxide and other polyatomic interferences were reduced by several orders of magnitude.

iv) As already discussed, one of the most effective methods of overcoming spectral interferences is to use high resolution ICP–MS, using a double focusing mass analyser to enable mass measurements at very high resolution (50,000 approx., the resolution is typically defined as $P = M/\Delta M$, where ΔM is the width of the peak at mass M at 5% of the peak height). At present these systems are finding applications in precise determination of stable isotopes and isotope ratios in geological and environmental studies[74].

v) One of the prominent areas of research in ICP–MS in recent times is the use of mixed gas plasmas[75] (addition of small amounts of oxygen or nitrogen or hydrogen to argon, or use of xenon or helium plasmas). The mixed gas plasmas have two major advantages in ICP–MS; reduction of polyatomic interferences and direct analysis of organic materials[76]. Use of helium as plasma gas is important because of its very high first ionization potential (24.5 eV) which exceeds that of argon (15.75 eV). Thus greater excitation and ionization of analyte ions is expected with the plasma. This is particularly important for sensitive estimations of elements such as halogens which are not efficiently ionized in argon plasma. Hence they are determined in negative ion mode in conventional ICP–MS.

Since its commercial introduction in 1983, ICP–MS has proved to be a versatile and powerful technique for handling routine analytical problems involving both elemental and isotopic analysis in a wide variety of matrices. By coupling alternative sampling techniques such as laser ablation, the ICP–MS technique is able to solve many analytical problems, not possible by other instrumental methods. Currently ICP–MS is a subject of very active research in the field of inorganic analytical chemistry and the instrument will continue to be an important asset to any chemical laboratory. The developments in ICP–MS, LA–ICP–MS and HR–ICP–MC–MS are expected to address both the technical and operational limitations of the present systems, eventually leading to more reliable, robust and low-cost instruments.

1. Ehrlich, H. L., *Geomicrobiology*, Marcel Dekker, New York, 1990, p. 646.
2. McLaren, J. W., in *The Handbook of Inductively Coupled Plasma Mass Spectrometry* (eds Jarvis, K. E., Gray, A. L. and Houk, R. S.), Blackie, USA, 1992, p. 265.
3. Emery, D. and Robinson, A., *Inorganic Geochemistry: Applications to Petroleum Geology*, Blackwell, London, 1993, p. 254.
4. Martin, M. H. and Coughtrey, P. J., *Biological Monitoring of Heavy Metal Polution* (ed. Mellanby, K.), Applied Science, London, 1992.
5. Rollinson, H. R., *Using Geochemical Data: Evaluation, Presentation, Interpretation*, Longman, England, 1993, p. 17.
6. Tolg, G., in *Metal Compounds in Environment and Life* (ed. Merian, E.), Science and Technology Letters, Middlesex, UK, 1992.
7. Hevesy, G. Von and Levi, H., *Nature*, 1936, **137**, 185.
8. Gordon, G. E., Randle, K., Goles, G. G., Corliss, J. B., Beeson, M. H. and Oxley, S. S., *Geochim. Cosmochim. Acta*, 1968, **32**, 369–396.
9. Gray, A. L., *Analyst*, 1975, **100**, 289–299.
10. Gray, A. L., *Dyn. Mass Spectrom.*, 1978, **5**, 106
11. Gray, A. L. and Date A. R., *Dyn. Mass Spectrom.*, 1981, **6**, 252–266.
12. Gray, A. L. and Date, A. R. *Analyst*, 1983, **108**, 159.
13. Houk, R. S., Fassel, V. A., Flesch, G. D., Svec, H. J., Gray, A. L. and Taylor, C. E., *Anal. Chem.*, 1980, **52**, 2283–2289.
14. Douglas, D. J. and French, J. B., *Spectrochim. Acta*, 1986, **B41**, 197–204.
15. Natarajan, V. and Rama Rao, V. V. K., Proceedings of the Sixth National Symposium on Mass Spectrometry, Dehradun, 1993, pp. 367–372.
16. Houk, R. S., *Anal. Chem.*, 1986, **58**, 97A–105A.
17. Moore, G. L., *Introduction to Inductively Coupled Plasma Atomic Emission Spectrometry*, Elsevier, Amsterdam, 1989, pp. 1–130.
18. Boumans, P. W. J. M., *Theory of Spectromechemical Excitation*, Hilger and Watts, London, 1966, Chapter 7.
19. Gray, A. L. F. Z., *Anal. Chem.*, 1986, **324**, 561–570.
20. Vaughan, M. A. and Horlick, G., *Appl. Spectrosc.*, 1986, **40**(4), 434–445.
21. Tan, S. H. and Horlick, G., *Appl. Spectrosc.*, 1986, **40**(4), 445–460.
22. Evans, E. H. and Giglio, J. J., *J. Anal. Atom. Spectrom.*, 1993, **8**, 1–18.
23. Burton, L. L. and Horlick, G., *Spectrochim. Acta*, 1992, **B47**, E1621–E1627.
24. Balaram, V., *J. Indian Chem. Soc.*, 1991, **68**, 600–603.
25. Balaram, V. and Saxena, V. K., Proceedings of the First International Conference on Plasma Source Mass Spectrometry, Durham, UK, 1988, abstract, p .44.
26. Balaram, V., Anjaiah, K. V. and Reddy, M. R. P., *Analyst*, 1995, **120**, 1401–1406.
27. Goossens, J., Vanhaecke, F., Moens, L. and Dams, R., *Anal. Chem. Acta*, 1993, **280**, 137–143.
28. Longerich, H. P., *At. Spectrosc.*, 1989, **10**, 112–115.
29. Balaram, V. *Geophys. Res. Bull.*, 1988, **26**, 1–7.
30. Balaram, V., Manikyamba, C., Ramesh, S. L and Anjaiah, K. V., *At. Spectrosc.*, 1992, **13**, 19–75.
31. Balaram, V., *At. Spectrosc.*, 1993, **14**, 174–179.
32. Fulford, J. E. and Quan, E. S. K., *Appl. Spectrosc.*, 1988. **42**, 3.
33. Jarvis, K. E., *J. Anal. At. Spectrom.*, 1989, **4**, 563–570.
34. Nagender Nath, B., Balaram, V., Sudhakar, M. and Pluger, W. L., *Marine Chem.*, 1992, **38**, 185–208.
35. Vanhoe, H., Vandecasteele, C., Versieck, J. and Dams, R., *Anal. Chem. Acta*, 1991, **244**, 259–267
36. Vijayalakshmi, S., Krishna Prabhu, R., Mahalingam, T. R. and Mathews, C. K., *At. Spectrosc.*, 1992, **13**, 61–66
37. Vanhecke, F., Vandocasteele, C. and Dams, R., *Anal. Lett.*, 1992, **25**, 919–936.
38. Kato, Takunori, Nakamura, S. and Morita, M., *Anal. Sci.*, 1990, **6**, 623–626.
39. Luck, J. and Siewers, U. F. Z., *Anal. Chem.*, 1988, **331**, 129–132.
40. Alswaidan, H. M., *Anal. Lett*, 1994, **27**, 145–152.
41. Paulsen, P. J., Beary, E. S., Bushee, D. S. and Moody, J. R., *Anal. Chem.*, 1988, **60**, 971–975.
42. Balaram, V., Ramesh, S. L. and Anjaiah, K. V. F. J., *Anal. Chem.*, 1995 (in press).
43. Balaram, V., Hussain, T. S. M., Uday Raj, B., Charan, S. N., Subba Rao, D. V., Anjaiah, K. V., Ramesh, S. L. and Ilangovan, I., *J. Geochem. Explor.*, 1995 (in press).
44. Beauchemin, D., *Trends. Anal. Chem.*, 1991, **10**, 71–77.
45. Ruzicka, J. and Fresenius, Z., *Anal. Chem.*, 1988, **329**, 653.
46. Plantz, M. R., Fritz, J. S., Smith, F. G. and Houk, R. S., *Anal. Chem.*, 1989, **61**, 149.
47. Denoyer, E. R., *At. Spectrosc.*, 1991, **12**, 215–224.
48. Whittaker, P. G., Lind, T., Williams, J. G. and Gray, A. L., *Analyst*, 1989, **114**, 675.
49. Gregoire, D. C., *Anal. Chem.*, 1990, **62**, 141.
50. Williams, J. G., Gray, A. L., Norman, P. and Ebdon, L., *J. Anal. At. Spectrom.*, 1987, **2**, 469–472.
51. Ebdon, L., Foulkes, M. E., Parry. H. G. M. and Tye, C. T., *J. Anal. At. Spectrom.*, 1988, **3**, 753.
52. Jarvis, K. E., *Chem. Geol.*, 1992, **95**, 73–84.

53. Huang, C. W. and Jiang, S. J., *J. Anal. At. Spectrom.*, 1993, **8**, 681–686.
54. Houk, R. S., *et al.*. *Anal. Chem.*, 1992, **64**, 2444
55. Walder, A. J., Platzer, I. and Fredman, P. A., *J. Anal. At. Spectrum.*, 1993, **8**, 19.
56. Powel, M. J., Quan, E. S. K., Boomer, D. W. and Wiederin, *Anal. Chem.*, 1992, **64**, 2253–2257.
57. Jakubowski, *et al.*, *Spectrochim. Acta*, 1992, **47**, 119–123.
58. Chen, *et al.*, *J. Anal. At. Spectrom.*, 1992, 7, 905–910.
59. Gray, A. L., *Analysist*, 1985, **111**, 551–556.
60. Jarvis, K. E., Gray, A. L. and Houk, R. S., *Handbook of Inductively Coupled Plasma Mass Spectrometry*. Blackie, Glasgow and London, 1992.
61. Denoyer, E. R., Fredeen, K. J. and Hager, J. W., *Anal. Chem.*, 1991, **63**, 445A–457A.
62. Pearce, N, J., Perkins, W. T., Abell, I., Duller, G. A. T. and Fuge, R., *J. Anal. At. Spectrom.*, 1992, **7**, 53–57.
63. Garcia Alonso, J. I., Babelot, J. F., Glatz, J. P., Cromboom, O. and Koch, L., *Radiochem. Acta*, 1993, in press.
64. Alonso, J. I., Garcia, Serrano, J. G., Babelot, J. F., Closset, J. C., Nicolaou, G. and Koch, L., personal communication, 1994.
65. Raith , A. and Hutton, R. C. Fresenius, *J. Anal. Chem.*, 1994, **349**, in press.
66. Walder, A. J. and Fredman, P. A., *J. Anal. At. Spectrom.*, 1992, 7, 571.
67. Walder, A. J., Koller, D., Reed, N. M., Hutton, R. C. and Fredman, P. A., *J. Anal. At. Spectrom.*, 1993, **8**, 1037–1041.
68. Halliday, A. N., Christensen, J. C., Hall, C. M., Jones, C. E., Lee, D. C., Teagle, D., Walder, A. J. and Entwistle, A., Proc. GEOANALYSIS 94: An International Symposium on the Analysis of Geological and Environmental Materials, Ambleside, England, 1994, C41.
69. Feldman, I., Tittes, W., Jakubowski, N., Stuewer, D. and Giessman, U., *J. Anal. At. Spectrom.*, 1994, **9**, 1007–1014.
70. Hall, G. E. M., *J. Geochem. Explor.*, 1982, **44**, 201–249.
71. Potter, D., *Int. Lab.*, 1994, **3**, 30–31.
72. Clifford, R. H., Tan, H., Liu, H., Montaser A., Zarrin, F. and Keady, P. B., *Spectrochim. Acta*, 1993, **B48**, 1221–1235.
73. Alves, L. C., Daniel, R., Wiederin, R. O. and Houk, R. *S.*, *Anal. Chem.*, 1992, **64**, 1164–1169.
74. Yamasaki, S. I., Tsumura, A. and Takaku, Y., *Microchem. J.* 1994, **49**, 305–318.
75. Durrant, S. F. F., *J. Anal. Chem.*, 1993, **347**, 389–392.
76. Hutton, R. C., *J. Anal. At. Spectrosc.*, 1986, **1**, 259–263.
77. Cousin. H. and Magyar, B., *Mikrochim. Acta*, 1994, **13**, 313–323.

ACKNOWLEDGEMENTS. I thank Dr S. M. Naqvi for his valuable guidance and encouragement; Dr Sajid Husain, Analytical Division, Indian Institute of Chemical Technology, Hyderabad, for critical reading of the manuscript and helpful suggestions; Drs S. Nirmal Charan and M. V. Subba Rao (NGRI) for their helpful suggestions.

Climate change and agriculture – An Indian perspective

Sulochana Gadgil

In this paper, we examine the major predictions made so far regarding the nature of climate change and its impacts on our region in the light of the known errors of the set of models and the observations over this century. The major predictions of the climate models about the impact of increased concentration of greenhouse gases are at variance with the observations over the Indian region during the last century characterized by such increases and global warming. It is important to note that as far as the Indian region is concerned, the impact of year-to-year variation of the monsoon will continue to be dominant over longer period changes even in the presence of global warming. Recent studies have also brought out the uncertainties in the yields simulated by crop models. It is suggested that a deeper understanding of the links between climate and agricultural productivity is essential for generating reliable predictions of impact of climate change. Such an insight is also required for identifying cropping patterns and management practices which are tailored for sustained maximum yield in the face of the vagaries of the monsoon.

THE last decade has witnessed a rapid increase in the awareness of global change and triggered widespread apprehension amongst scientists and governments about the implications for their part of the globe. The major-facets of global change are changes in climate, the concentration of atmospheric constituents (such as CO_2, CH_4, etc.), land surface cover (e.g. desertification, deforestation) and biodiversity. The changes in these facets are of concern to us because of the impact on human beings and on resources critical to their survival. On the other hand, the changes in many of these facets have a large anthropogenic component, e.g. deforestation, loss in biodiversity, the increase in CO_2 due to burning of

Sulochana Gadgil is in the Centre for Atmospheric Sciences, Indian Institute of Science, Bangalore 560 012, India.

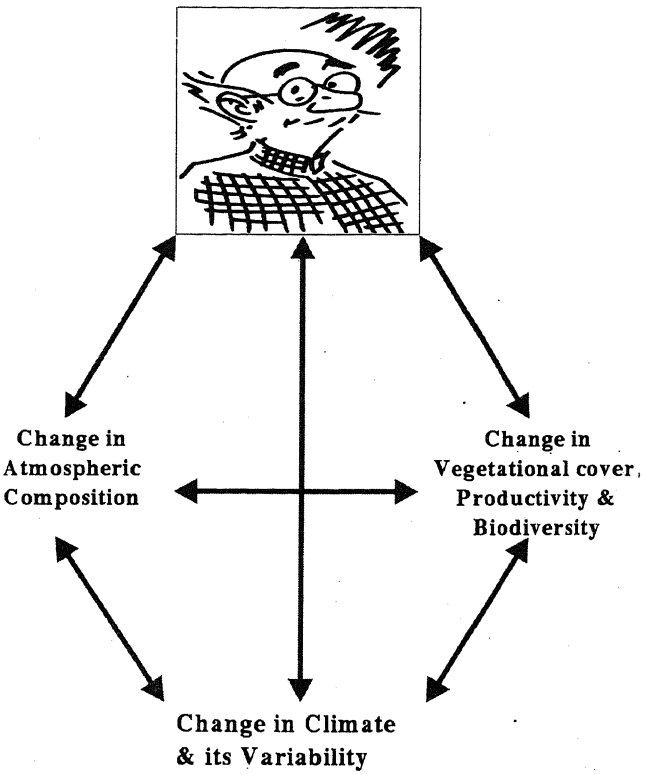


Figure 1. Major facets of global change – induced by and having an impact on humans.

fossil fuels, etc. It is important to note that many of these changes are interlinked (Figure 1). Thus increase in CO_2 (which is a greenhouse gas) can lead to climate change. Change in land surface cover can also have a large impact on the climate and particularly the rainfall in the tropics. Changes in atmospheric constituents can either have a direct impact on the inhabitants of the earth (such as higher incidence of skin cancer due to the depletion of ozone caused by the chloroflurocarbons (CFCs)) or an indirect impact via the induced change in climate as expected from the increase in CO_2.

I focus here on climate change. Climate change on long timescales of thousands of years, as between ice-ages and interglacials, is believed to be the result of natural factors such as changes in parameters of the earth's orbit around the sun. Anthropogenic factors are likely to have an impact on the timescales of about a century. Perhaps the most prominent facet of climate change over this century is global warming. The mean global surface temperature exhibits an increase over the last decades, with particularly sharp increase from 1970s (Figure 2). This warming has occurred during a period in which the composition of the atmosphere has changed significantly with a marked increase in the concentration of several gases contributing to the greenhouse effect such as carbon dioxide and methane (Figure 3). A major intergovernmental effort[1] in scientific assessment of climate change (IPCC) concluded that 'we are certain that emissions resulting from human activities are substantially increasing the atmospheric concentration of greenhouse gases such as CO_2, methane, CFCs, etc. These enhance the greenhouse effect, resulting in an additional warming of the earth's surface (on an average). The main greenhouse gas, water vapour, will increase in response to global warming and further enhance it.'

Several predictions have been made about the nature of climate change induced by an increased concentration of greenhouse gases and the possible impacts on critical resources such as agriculture. For example, a major study of implications for India of climate change[2] concludes 'cereal production is estimated to decrease and the nutrition security of the population-rich but land-hungry region of India would be hampered, an increase in the local tropical cyclone activity may occur over the next century, posing serious problems as large areas in the coastal regions have a dense population.'

Most of the studies using climate models to estimate change associated with enhancement of greenhouse gases suggest that the summer rainfall over South Asian region will increase by 5 to 15% (ref. 1). Some studies

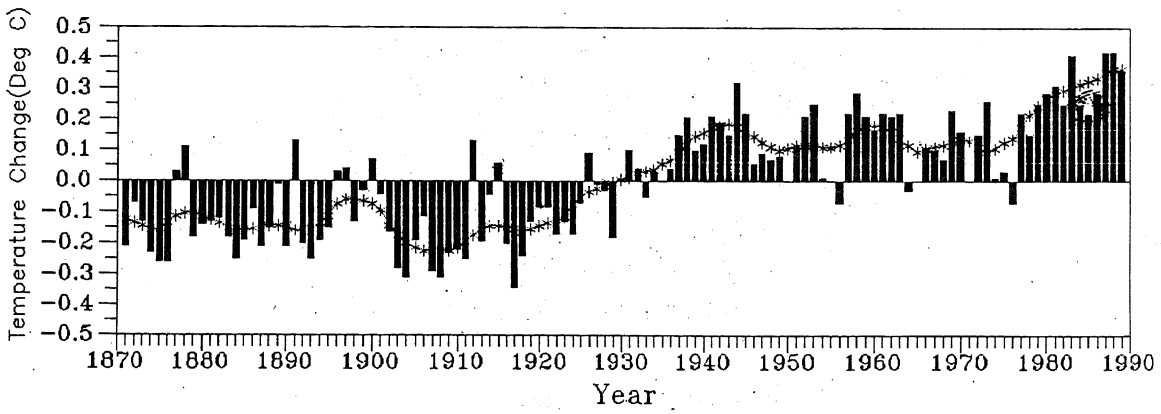


Figure 2. Annual deviation of the global mean combined land–air and sea-surface temperature for the period 1861–1989. The curve shows the results of a smoothing filter applied to the annual values.

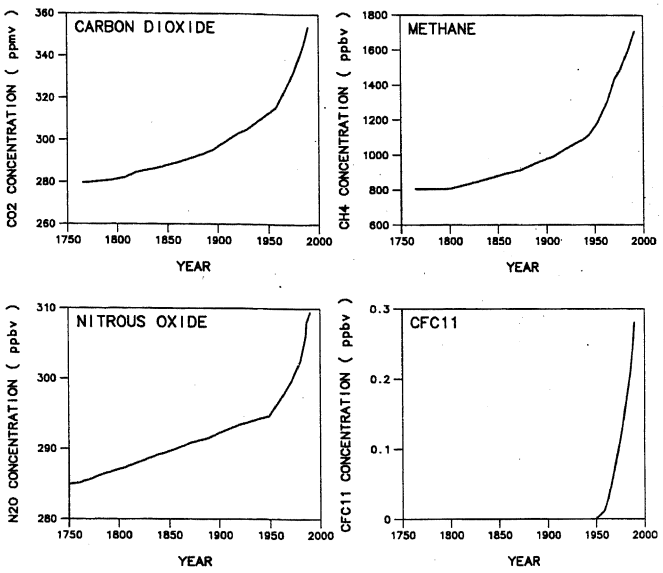


Figure 3. Variation of the concentration of some atmospheric constituents during the last century.

predict an increase in the intensity of daily rainfall events. In the thematic review of agriculture 'Climate Change over Asia', Suppiah *et al.*[3] state, 'Studies of simulated daily rainfall from enhanced greenhouse experiments indicate that a systematic increase in the average intensity is a common feature. Associated with this is an increase in the frequency of heavy rainfall events. Indeed, there is a reason to be more confident that rainfall intensity will increase in south and southeast Asian regions under enhanced greenhouse conditions. This pattern of change has serious implications for agricultural activities.' Parry and Swaminathan[4] suggest that 'the changes in climate could provoke serious reduction in food production potential, particularly in regions that are at present vulnerable to climate variability such as the agricultural zones in semi-arid and humid tropics'.

Such predictions about the nature of climate change and its impacts on agriculture are widely publicized. Thus the headlines of an article based on a study of Rosenzweig and Parry[5], which appeared in several newspapers across the country, read 'Impact of global warming to be hardest on have-nots. Less food if earth warms; more hunger in developing countries'. However, there is hardly any discussion about how reliable these predictions are. Predictions about the nature of climate change due to enhancing the concentration of greenhouse gases are generally made using climate models which are run for the present concentration and enhanced concentration of these gases. Assessments of impacts on agriculture such as that of Rosenzweig and Parry[5] involve using outputs of such models in conjunction with crop models and models of world trade. There are systematic errors in each of these models. Obviously, an assessment of the reliability of such predictions over the region of interest – here India, is an essential prerequisite to generating policy recommendations about possible mitigatory measures, and the strategies of research and development.

In this article, the major predictions made so far regarding the nature of climate change and its impacts for our region are examined in the light of the known errors of the set of models and observations over this century. The aim of this article is to identify problems which are of particular importance for our region and start a dialogue amongst concerned scientists to evolve a programme of research and development appropriate to the Indian situation.

Nature of climate change – Predictions and observation

Let us consider first the use of climate models for prediction of climate change and variability over the S. Asian monsoon region. It is important to note that simulation of the summer monsoon over this region has proved to be an extremely challenging problem. This is evident from Figure 4, in which the observed pattern of the mean rainfall for July is shown along with the mean pattern simulated by three atmospheric general circulation models (GCMs) for the period 1979–88 as a part of a

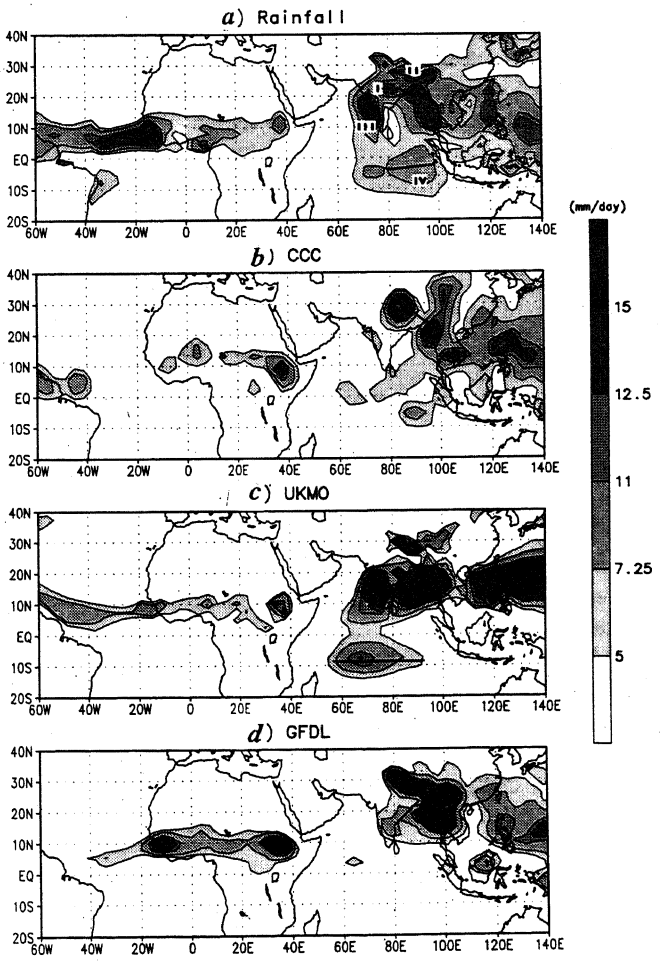


Figure 4. Mean rainfall pattern for July (*a*) observed patterns (*b*), (*c*), (*d*) simulated by three different atmospheric GCMs.

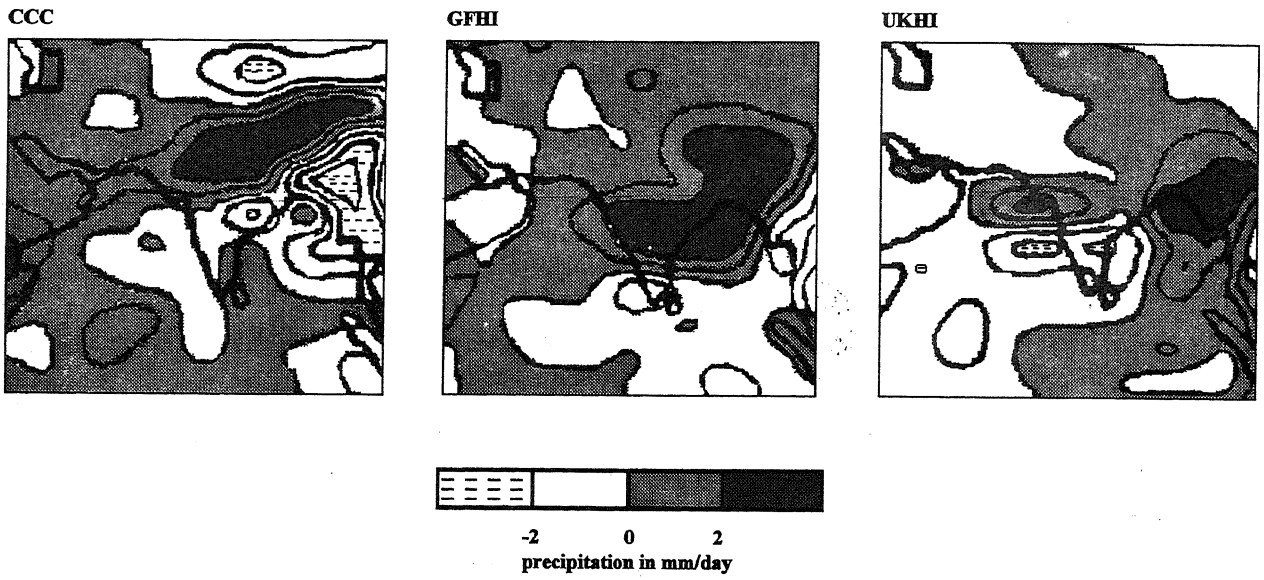


Figure 5. Change in precipitation (smoothed 10-year means) due to doubling CO_2 during June–August as simulated by some models (adapted from IPCC).

major Atmospheric Model Intercomparison Experiment[6]. It is important to note from Figure 4 *a* that while over the African region there is only one rainbelt around 10°N, over the Indian longitudes there are four. The large-scale monsoon rainfall is associated with the rainbelt I which extends north westward from the Bay of Bengal onto the Indian subcontinent. Northward of this monsoon rainbelt, is the rainbelt II along the foothills of the Himalayas. In addition, there is the rainbelt III along the west coast of the Indian peninsula due to the Sahyadris, and the rainbelt IV on the warm waters of the equatorial Indian ocean. The presence of so many favourable lacations for the rainbelt makes it extremely difficult for models to simulate the rainfall pattern realistically[7]. Generally only some of the rainbelts are simulated and that too not always at a realistic location. Thus, the CCC model simulates only the rainbelt II and the oceanic rainbelt IV. In the UK Meteorological Office (UKMO) simulation, the rainbelt over the continent is primarily on the peninsula, with hardly any rain over the monsoon zone (rainbelt I); the Himalayan belt II and oceanic belt IV are also simulated. The most realistic simulation of rainbelt I in Figure 4 is by the GFDL model. However, the oceanic belt IV is totally absent in the GFDL simulation. The simulation of the year-to-year variation of the summer monsoon over the Indian region is found to be even more difficult than that of the average pattern. Thus even models which simulate the mean pattern reasonably well, are not able to simulate realistically the interannual variation between poor and good monsoons[7].

So far we have been discussing simulation by atmospheric GCMs for the period 1979–88 with the sea surface temperature specified from observations. Because of the large inertia, the timescales of response of the ocean are much longer than those of the atmosphere. Hence for generating predictions over timescales of a week or less, or for studying subseasonal variation (such as fluctuations of the monsoon between active and dry spells), atmospheric GCMs run with specified sea surface temperatures are useful. However, for simulating variations on long timescales such as the century-scale on which we expect a change of greenhouse gas concentration to have an impact, it is necessary to use climate models in which the ocean is also an active player. These climate models are models of the coupled atmosphere–ocean system, which are necessarily more complex than the component atmospheric or ocean models. Not surprisingly, the simulation of the summer monsoon rainfall pattern by such models is also generally not realistic (e.g. Lal *et al.*[8]).

Some differences in the simulated and observed patterns of rainfall and its variability are to be expected. As Gates[9], one of the leaders in the field, has pointed out, 'The demand for climate scenarios for impact estimation far exceeds the supply of reliable information climate modelers have in their possession. Climate models certainly have systematic errors, most of these errors are poorly understood and, in general, poorly documented'. We have seen that it appears particularly difficult to simulate the South Asian monsoon and its variability by atmospheric GCMs or climate models, perhaps because of the presence of multiples equilibria[7]. Such limitations have to be borne in mind when we consider the implications of climate change derived by climate models.

Let us consider next the predictions made for enhanced greenhouse gases. The most authoritative review is by the IPCC[1]. While all models indicate that surface temperature

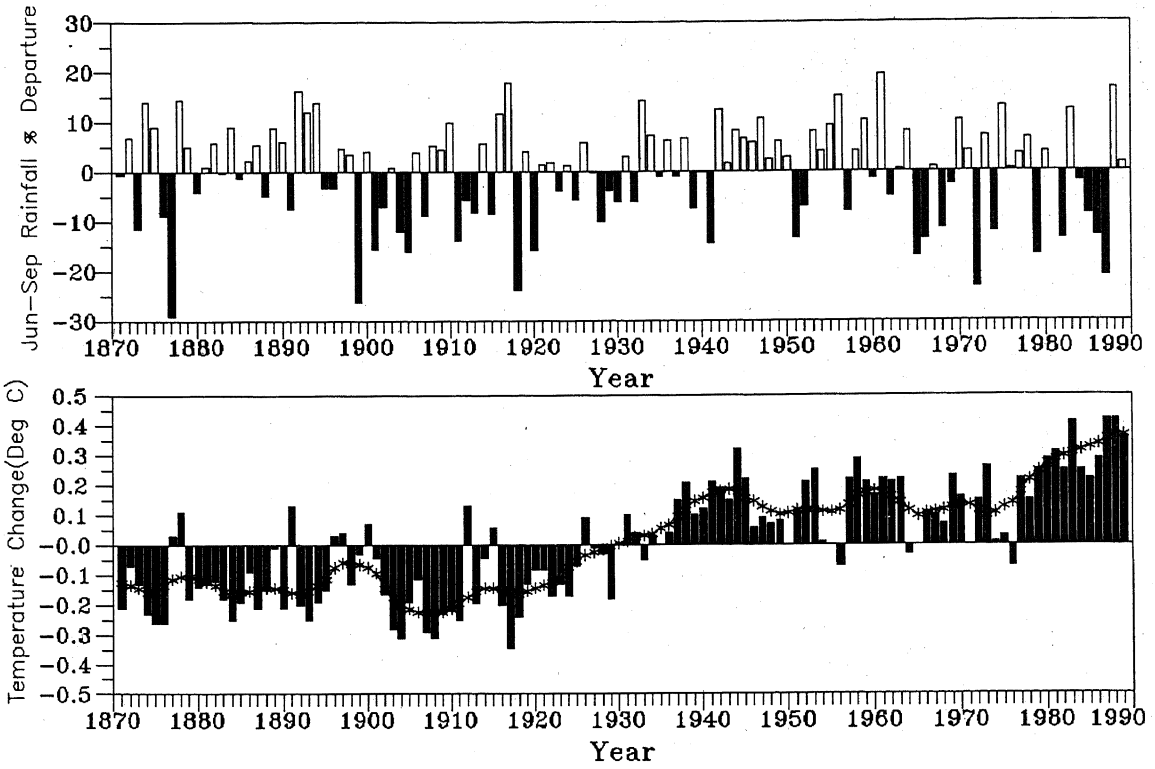


Figure 6 *a*. Variation of the all-India summer monsoon rainfall over the last century. Global warming of Figure 1 is also shown.

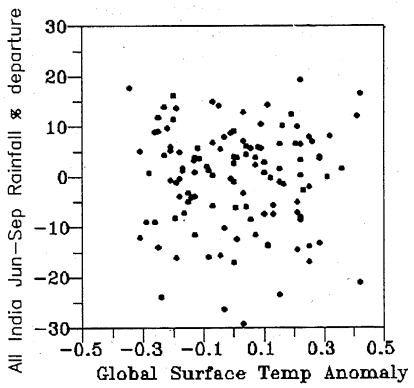


Figure 6 *b*. Percentage departure from normal of all-India summer monsoon rainfall versus global surface temperature anamoly.

will increase over the Indian region (up to 2°C during the summer) the response in terms of precipitation changes is more difficult to assess. Although most models simulate an enhancement of precipitation over the South Asian monsoon region, the change in precipitation is relatively small compared to the interannual variations. Also, there are considerable differences between models (Figure 5). However, all the three models shown in Figure 5 imply an increase over the region north of 20° N, i.e. over the monsoon zone of Figure 4 *a*, over which most of the large-scale monsoon rainfall is received. Lal *et al.*'s[10] study of anticipated changes in India's climate also suggests an increase in rainfall over this region.

Let us now consider how the prediction of an increase in the Indian monsoon rainfall in association with global warming compares with observations. The observed variation of the all-India summer monsoon rainfall over India over the last 100 years and the global warming during this period are shown in Figure 6 *a*. The variations of all-India rainfall over this short period of about a century do not seem to be correlated with the observed global warming (Figure 6 *b*). We note that there is no clear increasing trend observed on the century scale in Figure 6. *Rather, the most prominent variation is on the interannual scale (from year to year), between good and poor monsoons*. The change on the decadal or century scale is in the frequency of deficit monsoons or droughts. Joseph[11] was the first to point out that there are three distinct epochs in this series – the first up to about 1920, characterized by frequent droughts; the period 1930–64 by relatively less frequent droughts and the period 1965–87 again by frequent droughts. Thus what is important for the Indian region as far as change in climate is concerned, is the change (if any) in the frequency of extreme events such as droughts of floods, i.e.

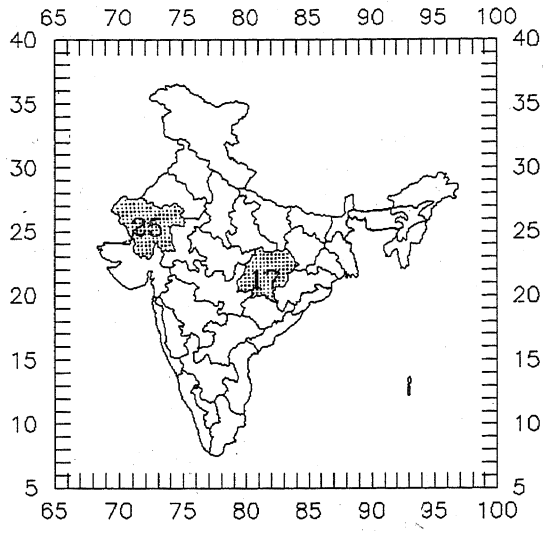


Figure 7 *a*. Coherent rainfall zones after Gadgil *et al.*[13].

changes in the important facets of climate variability. Although there are changes in the frequency of droughts on the decadal scale, these do not appear to be simply related to global warming[12]. For example, although the epoch of frequent droughts till 1920 coincides with the relatively low global temperatures, and the warming up to 1945 occurred in an epoch of low frequency of droughts, in the warm epoch during 1965–88, a high frequency of droughts is observed.

It is well known that the variations of seasonal rainfall over India do not occur in phase across the country. When some parts of India experience drought, there are floods in others, and so the average all-India rainfall is not very meaningful. Hence, we consider variations of rainfall over zones, which are coherent with respect to variations of the summer monsoon[13]. We find that there is no clear trend of increasing or decreasing rainfall over the century scale for any of the zones although there does appear to be some signal on decadal scales, leading to a persistence of above or below average rainfall for several years in a row (e.g. Figure 7). Earlier studies of trends over subdivisions or macrozones (reviewed by Pant *et al.*[14]) also found no significant trend over most of these regions. Rupa Kumar *et al.*'s[15] analysis of rainfall time-series at 306 stations across India showed that very few stations had significant trends.

An increase in the intensity of daily rainfall and frequency of heavy rainfall events has also been predicted[16]. The variation of the frequency of the heavy

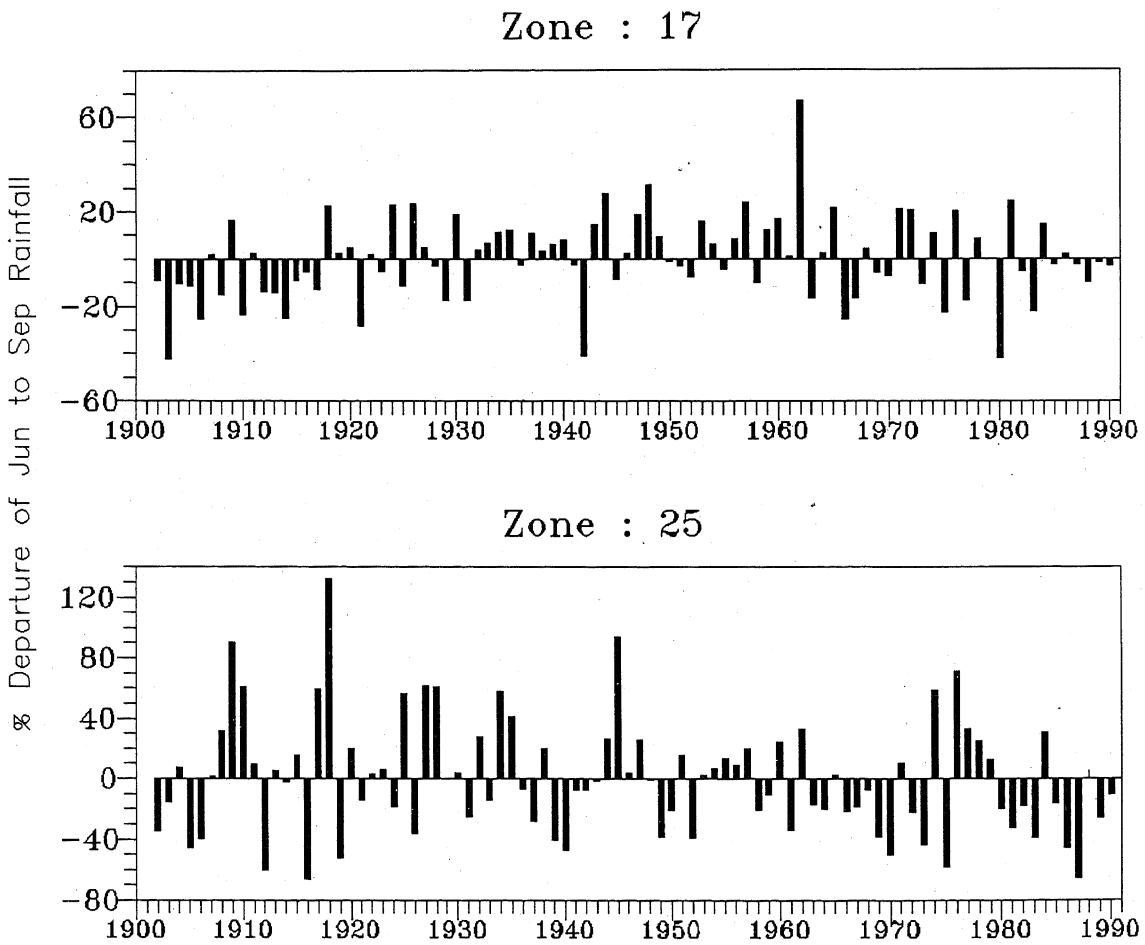


Figure 7 *b*. Variation of the summer monsoon rainfall over one zone 25 (below) over the North Western part of the monsoon zone, and zone 17 (above) over the eastern part of the monsoon zone.

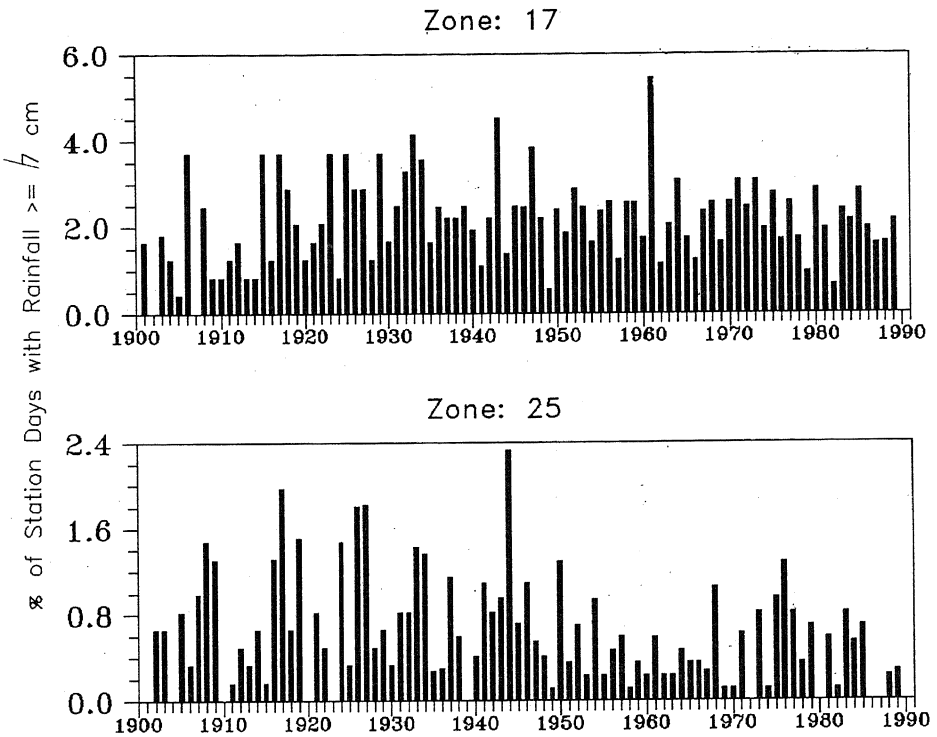


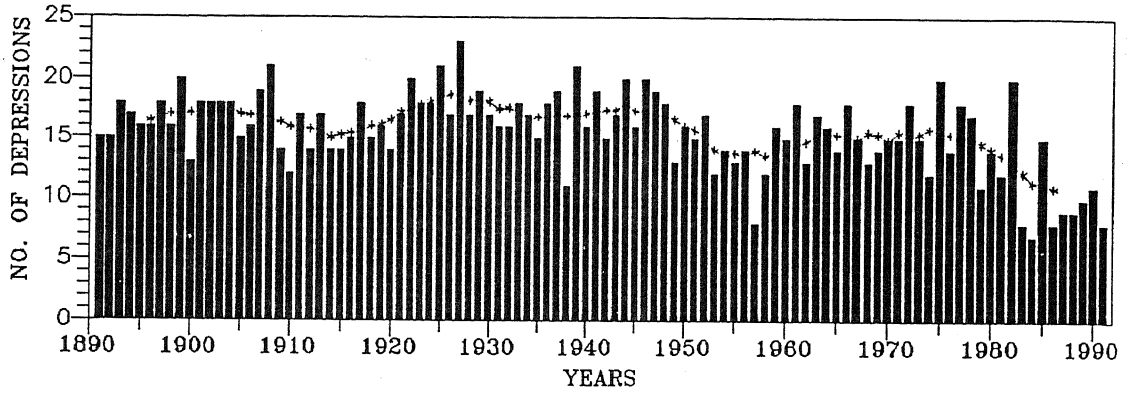
Figure 8. Variation of the frequency of heavy rainfall events (> 7 cm per day) over the two zones in Figure 7.

rainfall (greater than 7 cm a day) for two zones of Figure 7 is shown in Figure 8. Again there is no clear trend indicating increase of heavy rainfall events in association with global warming. In fact the frequency appears to decrease in the last two decades.

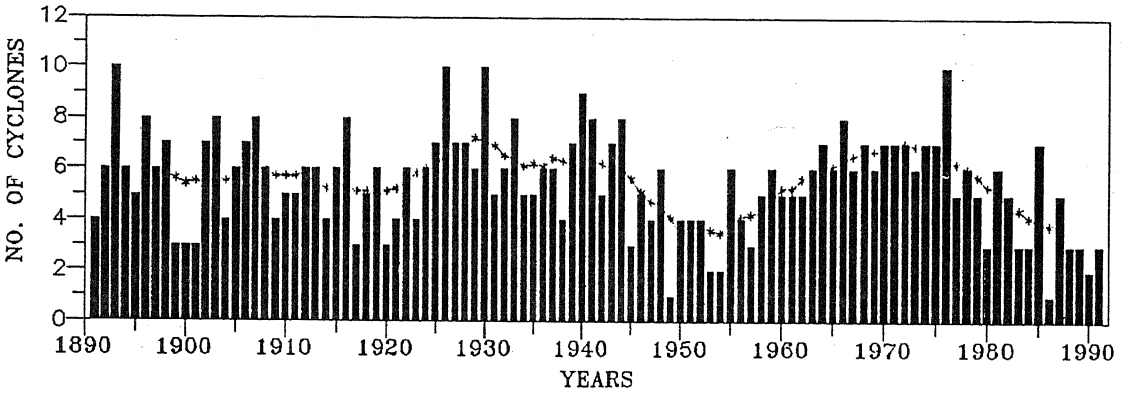
Another prediction is about a possible increase in frequency and intensity of tropical depressions and cyclones[10]. A detailed discussion about the nature of change in the depressions and cyclones over the last 100 years appears elsewhere[17]. The number of depressions and cyclones north of the equator in the Indian Ocean from Joseph's[17] study are shown in Figure 9. During the period 1970–90 when global warming is most rapid, the frequency of these systems is in fact decreasing. The frequency of tropical cyclones over the N. Atlantic also appears to have decreased in this period while there is no clear trend over the Pacific[18].

Studies of the observed changes in surface temperature over the Indian region have also been reviewed by Pant *et al.*[14]. The all-India mean annual surface temperature derived from 73 stations across India[19] shows a significant warming of 0.4°C per 100 years, which is comparable to the global mean trend of 0.3°C per 100 years. This increase is primarily contributed by the post-monsoon and winter seasons, there being hardly any trend in the monsoon-season temperatures. Of the 73 stations used to derive the all-India average, 30 stations show a significant warming trend while 6 show a cooling trend. *A comparison of the trends in rainfall[15] and temperature[19] indicates that generally increase in rainfall is associated with decrease in temperature.* In the belt north of 20°N where most models predict an increase, the eastern sector exhibits a decreasing trend while the western an increasing one, although neither is significant at a majority of the stations.

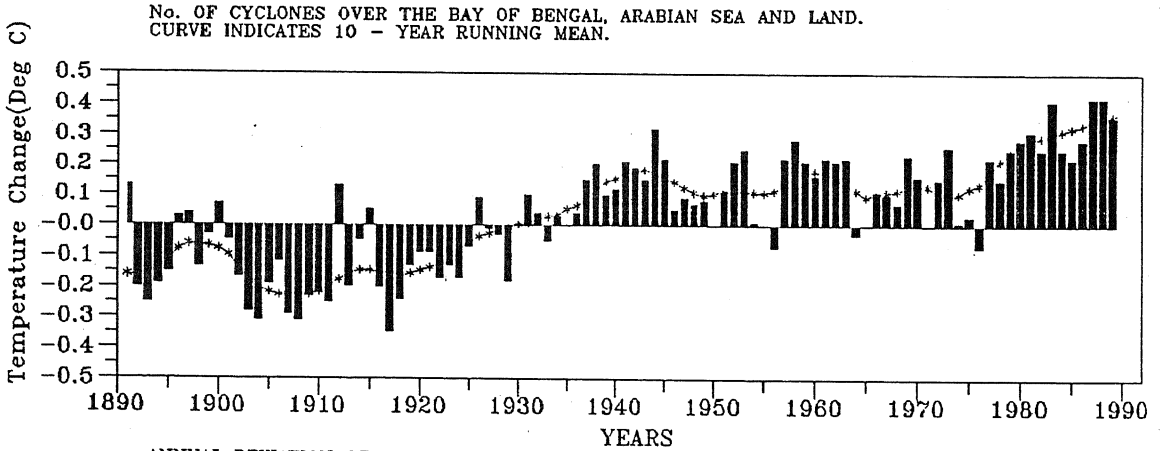
The observations about the variation of the summer monsoon rainfall over the Indian region over the last century during the period of global warming can be summarized as follows. The all-India summer monsoon rainfall does not exhibit an increasing trend; its variations do not appear to be correlated with global warming. No trends are apparent in the variation of the summer monsoon rainfall over different coherent zones of Indian region. There is no evidence to suggest that an increase in heavy rainfall events occurs with global warming. The frequency of depressions and cyclones over the Indian seas appears to have decreased markedly over the period 1970–90 when global warming was most



No. OF DEPRESSIONS OVER THE BAY OF BENGAL, ARABIAN SEA AND LAND.
CURVE INDICATES 10 - YEAR RUNNING MEAN.



No. OF CYCLONES OVER THE BAY OF BENGAL, ARABIAN SEA AND LAND.
CURVE INDICATES 10 - YEAR RUNNING MEAN.



ANNUAL DEVIATION OF GLOBAL MEAN COMBINED LAND-AIR AND SEA-SURFACE TEMPERATUREES
FOR THE PERIOD 1891-1989. THE CURVE SHOWS THE RESULTS OF A SMOOTHING FILTER
APPLIED TO THE ANNUAL VALUES

Figure 9. Variation of the frequency of depressions and cyclones over the Indian seas.

rapid. Hence the observations about the changes in rainfall pattern over the Indian region during this century appear to be at variance with the predictions generated so far by the existing climate models.

The all-India surface temperature does exhibit a warming as predicted by the models, albeit at a much smaller rate. However, over regions where rainfall increase is observed (again of very small amplitude) generally the temperature decreases. So the observations suggest that a simultaneous increase in rainfall and temperature is unlikely.

This comparison of model predictions regarding the important climatic variables – temperature and rainfall with available observational data – suggests that until the performance of the models in simulation for the regional scale is sufficiently accurate, regional predictions of the impact of various factors such as enhancement of greenhouse gases generated on the basis of the results from such models have to be interpreted with caution. The observations may not be adequate for detecting the climate change signal from the natural variability. However, before making decisions regarding actions for

mitigation, validation of the models by comparison with whatever observations are available for the major predictions, is absolutely essential.

Climate and agriculture

The strong links between the vagaries of the monsoon and agricultural productivity are well known. This link is evident in the variation of the production of rice and of the annual production of total food grain over India, shown in Figure 10, along with the all-India summer monsoon rainfall. It is clearly seen that superposed on the increasing trend due to the green revolution, are dips in the yield, which occur in years in which the all-India summer monsoon rainfall is deficit (e.g. 1965, 66, 72, 74, 79, 82). On the other hand, when the all-India rainfall is in excess of the normal (e.g. 1964, 70, 75, 78, 83), the yield is certainly higher. Thus the large year-to-year variations in monsoon rainfall are directly reflected

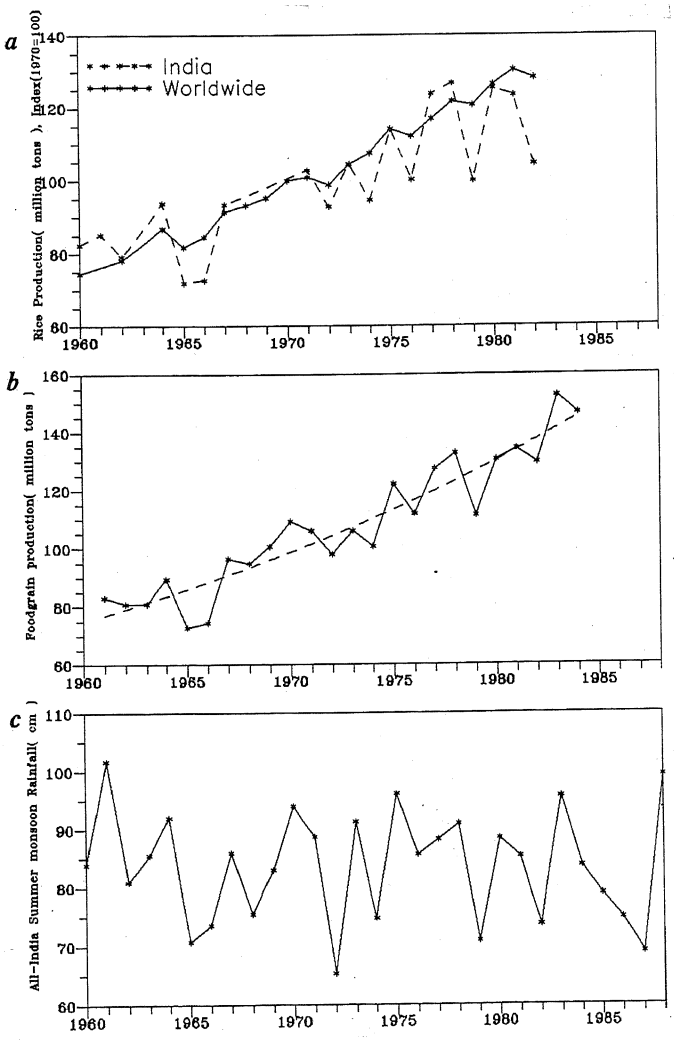


Figure 10. Variation during 1960–82 of (*a*) rice yield over the Indian region (after Swaminathan[23]), (*b*) total food grain production (after Parthsarathy *et al.*[24].), (*c*) all-India rainfall.

in substantive variations of the yield. For example, the difference in the total annual food grain production between the poor monsoon of 1974 and the succeeding good monsoon of 1975 is about 20% of the yield.

We note that in the epoch 1930–64 when the mean monsoon rainfall was high, the frequency of deficit years was low. Thus *a priori*, it would appear that if the monsoon rainfall were to increase (and frequency of droughts to decrease) as predicted by the climate models, the production of food grains would be enhanced. What then is the basis of the sensational predictions about decrease in yield with global warming? These are generated by the use of crop models to assess the impact of change in climate.

In general, the change in yield depends not only on the change in rainfall but also change in CO_2 concentration and temperature. While the increase in rainfall and concentration of CO_2 generally lead to an increase in yield, the impact of an increase in temperature is generally negative in the tropics. Thus if both rainfall and temperature increase, it is necessary to ascertain whether the impact of increased rainfall and CO_2 will overwhelm the impact of increased temperature and hence evapotranspiration over specific regions for specific crops. However, as noted in the last section, it is not clear that over the Indian region temperature and rainfall will simultaneously increase.

So far we have focused on interannual variation and considered primarily the impact of fluctuations in the total seasonal rainfall, mean temperature, etc. However, the agricultural productivity is also a sensitive function of the distribution of the meteorological variables such as rainfall within a season[20]. The difference between a so-called good monsoon season and a deficit monsoon season is generally the presence of a long dry spell in the latter. Such dry spells induce moisture stress and have a major impact on the productivity when they occur in critical life history stages of the crop. Similarly if wet spells occur and promote the spread of pests and diseases, significant losses are incurred unless mitigatory measures are adopted[21]. The identification of such weather/climate events which have a large impact on the productivity of different crops is essential for making realistic impact assessments and choosing the appropriate management strategies for mitigatory measures.

With the rich data-set available in the meteorology department, the probabilities of the occurrence of such events in different stages in crop growth can be readily worked out for each region. For obtaining maximum sustainable yield in the face of this variability, the varieties and cropping patterns have to be chosen so that the occurrence of such events in the critical life-history stages is minimized. Thus, understanding such links between climate variability and agricultural productivity is also a pre-requisite to tailoring cropping patterns/varieties and management practices for maximum sustainable yields in different agroclimatic zones.

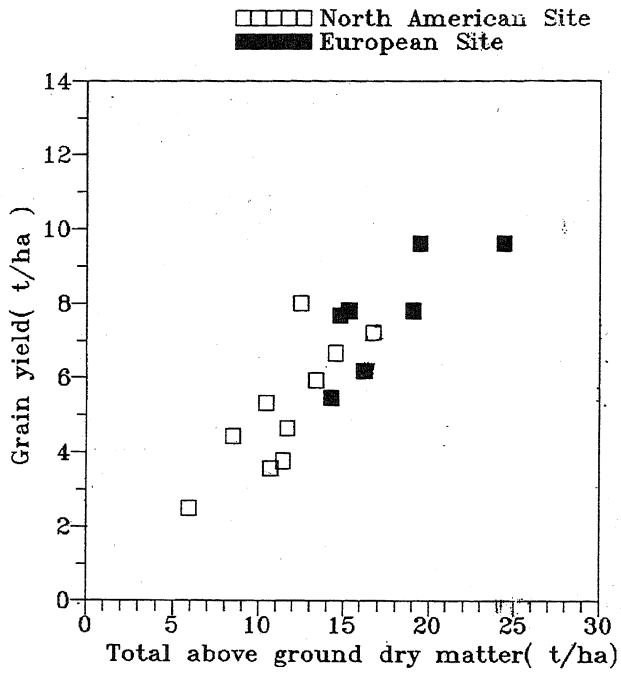


Figure 11. Scatter plot of simulated yield against biomass for the North American site and European site in the GCTE (after Ingram[22]).

A systematic analysis of the yields simulated by wheat models was recently conducted as a part of the IGBP Project on Global Change and Terrestrial Ecosystems (GCTE). This revealed a surprisingly large intermodel variation in simulated yield even when the climatic conditions were specified to be the same (Figure 11, after Ingram[22]). Thus it appears that the prediction of impact generated by using crop models in combination with climate models are likely to have significant errors. It is necessary to understand and validate each component/link with available observations and special field trials to determine the reliability of the predictions. For example, it may be worthwhile to investigate whether the existing crop models (of the kind used in Rosenzeweig and Parry's[5] study) can simulate the observed variation of the yields over the last century when the observed variation of the climatic parameters is specified.

It is also important to note that as far as the Indian region is concerned, the impact of year-to-year variations of the monsoon will continue to be dominant even in the presence of global warming. Thus neither the rainfall nor temperature is likely to change monotonically. The pertinent question to ask of climate models and observations is then the extent to which climate variability will change. In other words, on the interannual scale whether the frequency of extreme events such as droughts or floods will change and on the subseasonal scale whether the frequency of critical weather events such as dry and wet spells or hot and cold spells will change.

Concluding remarks

We have seen that predictions generated from climate models about the consequences of enhanced greenhouse gases such as increased monsoon rainfall, higher frequency of heavy rainfall events, etc. are at variance with observations over the last century characterized by global warming. Recent studies have also brought out the uncertainties in the yield simulated by crop models. Thus it is clear that understanding the important links between climate and agricultural productivity is an essential prerequisite to generating reliable predictions about impact of change in climate and variability.

In my view, the most important problem for our region is to elucidate in detail the climate component of agricultural productivity by systematic analysis of the available data on yields of different crops (at the different agricultural research stations as well as at district level) on the one hand and investigation and development (if necessary) of crop models with realistic sensitivity to climate on the other. This will pave the way for generating realistic impact assessments and identifying cropping patterns and management strategies which are tailored for sustained maximum yield in the face of the vagaries of the monsoon.

Meanwhile, we expect the rapid improvement of climate models across the globe in the last decade to be sustained. In the not too distant a future, these models should start yielding reliable results on regional scale for the nature of climate change in response to various factors such as enhanced greenhouse gases. By that time if the links to the productivity of the major crops are understood, reliable predictions of impact could be generated and strategies for coping with the anticipated climate change worked out.

1. Houghton, J. T., Jenkins, G. J. and Ephraums, J. J., *Climate Change: The IPCC Scientific Assessment*, World Meteorological Organization, 1990.
2. Govinda Rao, P., Kelly, M., Hulme, M. and Srinivasan, G., Centre for Social and Economic Research on the Global Environment, University of E. Anglia, UK, Working Paper GEC 94–22, 1994.
3. Suppiah, R., Pittock, B. and Whetton, P., in *Climate Change in Asia: Thematic Overview*, Asian Development Bank, 1994, pp. 16–56.
4. Parry, M. L. and Swaminathan, M. S., in *Confronting Climate Change: Risks, Implications and Responses* (ed. Mintzer, I. M.), Cambridge University Press, Cambridge, 1992, pp. 113–125.
5. Rosenzweig, C. and Parry, M. L., *Nature*, 1994, **367**, 133–138.
6. Gates, W. L., *Bull. Am. Meteorol. Soc.*, 1992, **73**, 1962–1970.
7. Gadgil, Sulochana and Raju, A., Proceedings of the first AMIP Workshop, Monterey, California, 15–19 May, 1995.
8. Lal, M., Cubasch, U. and Sanker, B. D., *Curr. Sci.*, 1994, **66**, 430–438.
9. Gates, W. L., in *Global Climate Change and California* (eds Knox, J. B. and Scheuring, A. F.), Univ. of California Press, California, 1991, pp. 58–68.
10. Lal, M., Bhaskaran, B. and Chakraborty, B., in *Global Warming* (ed. Lal, M.), Tata McGraw Hill, New Delhi, 1993, pp. 117–135.

11. Joseph, P. V., Proceedings of the Symposium on Tropical Monsoons, IITM, Pune, 1976, pp. 378–387.
12. Sikka, D. R. and Pant, G. B., Proceedings of Indo-US Workshop on Impact of Global Climate Changes on Photosynthesis and Plant Productivity, New Delhi, 1991, pp. 551–572.
13. Gadgil, Sulochana, Yadumani and Joshi, N. V., *Int. J. Climatol.*, 1993, **13**, 547–566.
14. Pant, G. B., Rupa Kumar, K. and Parthasarathy, B., in *Global Warming* (ed. Lal, M.), Tata McGraw Hill, New Delhi, 1993, pp. 71–91.
15. Rupa Kumar, K., Pant, G. B., Parthasarathy, B. and Sontakke, N. A., *Int. J. Climatol.*, 1992, 12.
16. Gordon, H. B., Whetton, P. H., Pittock, A. B., Fowler, A. M. and Haylock, M. R., *Climate Dynam.*, 1992, **8**, 83–102.
17. Joseph, P. V., 1995, Changes in frequency of depressions and cyclones over the Indian seas during the last hundred years (submitted).
18. Lighthill, J., Holland, G., Gray, W., Landsea, C., Craig, G., Evans, J., Kurihara, Y. and Guard, C., *Bull. Am. Meteorol. Soc.*, 1994, **74**, 2147–2157.
19. Hingane, L. S., Rupa Kumar, K. and Ramana Murthy, Bh. V., *Indian. J. Climatol.*, 1985, **5**, 521–528.
20. Jodha, N. S., Virmani, S. M., Huda, A. K. S., Gadgil, S. and Singh, R. P., in *The Impact of Climatic Variations on Agriculture* (eds Parry, M. L. *et al.*), Kluwer, Dordrecht, 1987.
21. Gadgil, Sulochana, Sheshagiri Rao, P. R., Joshi, N. V. and Sridhar, S., *Curr. Sci.*, 1995, **68**, 301–309.
22. Ingram, J., *Global Change: News Lett.*, 1994, No. 18, pp. 4–5.
23. Swaminathan, M. S., in *Monsoons* (eds Fein, J. S. and Stephens, P. L.), Wiley, 1987, pp. 121–133.
24. Parthasarathy, B., Munot, A. A. and Kothawale, D. R., *Agric. For. Meteorol.*, 1988, **42**, 167–182.

ACKNOWLEDGEMENTS. I am grateful to Sri R. K. Laxman for giving me permission to use his 'common man' in Figure 1, and to Prof. J. Srinivasan for discussions and suggestions for modification of the paper.

REVIEW ARTICLES

Opioid peptides in invertebrates: Localization, distribution and possible functional roles

R. Nagabhushanam, R. Sarojini, P. S. Reddy, M. Devi and M. Fingerman

Department of Ecology, Evolution, and Organismal Biology, Tulane University, New Orleans, Louisiana 70118, USA

In this review the opioids that have been found in the invertebrates are discussed. Major emphasis is placed on the arthropods, molluscs and annelids. The article deals largely with the identification and functional roles of these peptides. These endogenous opioids are involved in a wide variety of physiological processes, often serving as neurotransmitter substances. Among the putative roles of opioids in invertebrates are mediation of the release of neurohormones, initiation of proper behavioural responses, control of thermoregulatory activities and mediation of immune responses. This review concludes with a section that suggests areas of future investigation, along with possible practical applications of these compounds.

IN recent years a number of vertebrate-type neuropeptides have been discovered in the invertebrates. The discovery of these vertebrate-type peptides in invertebrate tissues has influenced research in comparative endocrinology in a major way. A class of such neuropeptides is known as the opioid family, based on the fact that among vertebrates these peptides show a potent capacity to mimic some key actions of morphine[1]. In mammals the enkephalins originate from at least two precursors, preproenkephalins A and B. Preproenkephalin A gives rise to one copy of leucine enkephalin (Leu-Enk), four copies of methionine enkephalin (Met-Enk) (Leu-Enk and Met-Enk are pentapeptides) and two C-terminally extended enkephalins, Met-Enk-Arg^6-Phe^7 (met-7) and Met-Enk-Arg^6-Gly^7-Leu^8 (Met-8)[2]. Preproenkephalin B contains no Met-Enk sequence but has three Leu-Enk-containing peptides, α/β-neoendorphin, dynorphin and rimorphin. The first enkephalins discovered, Met-Enk and Leu-Enk, were originally purified from the porcine brain[3]. There are a variety of opiate receptors (mu, delta, iota, kappa, epsilon) each of which appears to be the preferential ligand for one of the endogenous opioid peptides in mammals[4]. Endogenous opioids have been detected in all vertebrate species investigated[5,6]. These substances serve widely throughout the body as neuromodulatory substances.

The search for the presence of endogenous opioids in invertebrates initially met with some uncertainty. Early reports did not support the presence in several invertebrates of opiate-receptor-binding sites[7,8]. However, the development of appropriate techniques, especially immunocytochemistry, radioimmunoassay (RIA) and high-performance liquid chromatography (HPLC), made it possible to discover opioid peptides in several invertebrates. Among the invertebrates, studies have been undertaken mainly with arthropods, molluscs and annelids, and in this review we emphasize the studies that have been done on the localization, distribution and possible roles of the endogenous opioids in these groups.

Crustacea

Localization and distribution

The first opioid study on a crustacean was reported in 1981 by Mancillas *et al.*[9]. Through immunocytochemistry they found Leu-Enk-like immunoreactivity in all the retinular cells of the spiny lobster *Panulirus interruptus* and the red swamp crayfish *Procambarus clarkii*. In addition, such immunoreactivity was also apparent in nerve fibres in chiasm 3 that run from the medulla interna to the medulla terminalis. Since then immunocytochemical mapping studies, which not only detect the presence of opioid peptides but also give their distribution, have been reported in a variety of crustaceans. Table 1 summarizes the peptides detected so far in crustaceans and the techniques used to demonstrate their presence[9–11,14–22]. Marino *et al.*[12] and Piccoli *et al.*[13], by using a receptor-binding assay, reported the presence of opioid-like peptides in the suboesophageal ganglion and brain of *Squilla mantis* and *Carcinus maenas*. There are a few reports about the presence of enzymes that degrade the opioid peptides in Crustacea. Colleti-Previero *et al.*[23] found three peptidases in the haemolymph of the crayfish *Astacus fluviatilis* that rapidly degrade Leu-Enk. Watabe *et al.*[24] reported evidence for the presence of β-endorphin-degrading enzymes in the cells of the hepatopancreas of the crayfish *Procambarus clarkii*.

Functions

Behaviour. The first study of a possible role of an opioid in any crustacean was that by Maldonado and Miralto[25] with *Squilla mantis*. They found that the defensive response, a rapid flexure of the abdomen, which can be experimentally induced by electric shock, can be modified by morphine. This drug increases the threshold current required to elicit the response. The loss of sensitivity is dose-related. Naloxone blocks the action of this opioid. Naloxone is a competitive opioid antagonist, binding to opioid receptors. Martinez *et al.*[26] found that when the crab *Gecarcinus lateralis* was placed into the activity-monitoring chambers, the locomotor activity increased. FK 33 824, a stable Met-Enk analogue, markedly enhances this initial activity but naloxone blocks this excitatory action of the opioid. In studies with another crab, *Chasmagnathus granulatus*, Lozada *et al.*[27] found that this crab assumes a defensive posture, with both chelae extended and the body elevated when an electric shock of 50 Hz, 1 s duration and at least 8 V is given. Morphine produces a dose-dependent reduction in the sensitivity of this crab to the electric shock. But when naloxone was coinjected with morphine, no reduction in sensitivity occurred, which suggests that the morphine was acting through opioid receptors and that such receptors are indeed involved in this reduction of sensitivity to the electric shock. In other behavioural studies with this crab[28,29], it was found that a danger stimulus in the form of a passing shadow elicits an escape response that habituates after repeated stimulation, and that this habituation appears to be mediated by endorphins. Support for this suggestion that habituation is the result of endorphin release is the observation that naloxone increases the responsiveness to the danger signal and after a habituation session which reduces the response level to the danger signal there is also an analgesic effect on the defensive response of this crab to an electrical shock. Very recently, Godoy and Maldonado[30] reported that the synthetic opioid analogue (D-Ala^2) Met-enkephalin (DAME) significantly reduces the escape response to a danger stimulus in the crab *Chasmagnathus granulatus* when administered within a dose range of 0.01–1.0 μg/g. A 0.1 μg/g dose of naloxone has no effect *per se* on the response, but when it is administered together with DAME, it completely blocks the decremental effect of this drug.

Pigmentary effectors. Crustaceans have two types of pigmentary effectors: (a) the chromatophores, which are responsible for colour changes and (b) the retinal pigments, which control the amount of light impinging on the rhabdom. Translocation of the pigments in the chromatophores and at least in the distal and reflecting pigments is clearly regulated by neurohormones.

(a) *Chromatophores.* The physiology of crustacean chromatophores has been reviewed recently[31]. In several species dual control of chromatophores by pigment-dispersing and pigment-concentrating neurohormones has been demonstrated. In this laboratory we have been particularly interested in identifying the neuroregulators that control the release of crustacean neurohormones. In the fiddler crab *Uca pugilator*, 5-hydroxytryptamine (5-HT) stimulates the release of red-pigment-dispersing hormone while norepinephrine (NE) stimulates the release of black-pigment-dispersing hormone. Met-Enk, but not Leu-Enk, was found to stimulate the concentration of the black and red pigments in intact *Uca pugilator*[32]. In *in vitro* experiments Met-Enk stimulated the release of black- and red-pigment-concentrating hormones from isolated eyestalks[32,33]. Naloxone blocked the action of Met-Enk in intact crabs and on isolated eyestalks. Met-Enk, like the classical aminergic neurotransmitters, has no direct effect on these chromatophores, as evidenced by the lack of effect on the chromatophores in isolated legs. Presumably, therefore, 5-HT, NE and Met-Enk exert their effect only indirectly, by stimulating the release of the appropriate specific chromatophorotropic neurohormones. However, surprisingly, β-endorphin was found to produce black-pigment dispersion in both intact *Uca pugilator* and in isolated legs[32]. Additional evidence for the involvement of enkephalinergic substances

Table 1. Identification and localization of opioid-like peptides in crustaceans

Peptide	Species	Localization	Techniques	References
Leu-Enk-like	*Cancer pagurus*	PO	Immunocytochemistry	15
	Carcinus maenas	SG, LG, ME, MI, MT, PO, segmental nerve connections to the thoracic ganglia, thoracic ganglia	Immunocytochemistry HPLC	10, 14, 17, 18
	Gecarcinus lateralis	Brain	RIA, HPLC	19
	Maja squinado	PO	Immunocytochemistry	15
	Panulirus interruptus	Retinular cells, nerve fibres in chiasm 3	Immunocytochemistry	9
	Portunus puber	PO	Immunocytochemistry	15
	Procambarus clarkii	Retinular cells, nerve fibres in chiasm 3	Immunocytochemistry	9
	Uca pugilator	Retinular cells, LG, SG, OP, MT, ovary, testis	Immunocytochemistry	11, 21
Met-Enk-like	*Carcinus maenas*	SG, hepatopancreas thoracic ganglia	HPLC, immunoflurescence, RIA, sequence analysis	10, 16–18
	Gecarcinus lateralis	Eyestalk, brain	HPLC, RIA	19
	Uca pugilator	Retinular cells, SG, LG, ,OP, MT, x-organ of MT, ovary, testis	Immunocytochemistry	11, 21
Met-Enk-Arg^6-Phe^7	*Carcinus maenas*	SG	HPLC	10
Met-Enk-Arg-Phe-amide (YGGFMRFamide)	*Homarus americanus*	Nervous system	Immunocytochemistry	20
β-endorphin-like	*Procambarus clarkii*	Hepatopancreas	RIA, Immunocytochemistry	22

HPLC: high-performance liquid chromatography; LG: lamina ganglionaris; MT: medulla interna; ME: medulla externa; MT: medulla terminalis; OP: optic peduncle; PO: pericardial organ; SG: sinus gland; RIA: radioimmunoassay.

in crustacean colour changes was provided by Martinez *et al.*[34], who found that although injection of the stable Met-Enk analogue FK 33 824 alone into *Gecarcinus lateralis* has no effect on the chromatophores, coinjection of this analogue with eyestalk extract strongly potentiates the black-pigment-dispersing and red-pigment-concentrating actions of the eyestalk extract. This potentiation is blocked by naloxone.

(b) *Distal retinal pigment.* NE produces light adaptation of the distal retinal pigment of *Uca pugilator* whereas dopamine (DA) induces dark adaptation[35,36], presumably by stimulating the release of the light-adapting and dark-adapting neurohormones, respectively. Met-Enk, like DA, produces dark adaptations of this pigment, but Leu-Enk has no effect on this pigment[37].

Blood glucose. The crustacean hyperglycaemic hormone (CHH) is found in the sinus gland. In several crustaceans CHH release has been found to be triggered by 5-HT (ref. 38). Recently, Lüschen *et al.*[17] and Rothe *et al.*[18] provided evidence that synthetic Leu-Enk, endogenous sequenced Leu-Enk isolated from thoracic ganglia of *Carcinus maenas* and purified Leu-Enk-like material from sinus glands of this crab, when injected into intact *Uca pugilator*, the test organism for these bioassays, produce a decrease in the haemolymph glucose concentration. Also, in CHH bioassay done again by injecting into intact *Uca pugilator*, it was found that eyestalks of *Carcinus* incubated in the presence of synthetic Leu-Enk release less than the basal level of CHH was compared to eyestalks incubated in saline alone. Furthermore, injection of naloxone into intact *Uca pugilator* prior to injection of Leu-Enk blocks the hypoglycaemic action of this opioid in this crab. In contrast, Leu-Enk does not affect the haemolymph glucose concentration in eyestalkless fiddler crabs, which is consistent with the hypothesis that the enkephalin acts by reducing CHH output from the sinus glands in the eyestalks. In the crayfish *Astacus leptodactylus*, Met-Enk decreases the haemolymph glucose level as well as the nocturnal peak in the circadian rhythm of haemolymph glucose[39]. Sarojini *et al.*[40] provided evidence that both L-Enk and DA are involved in the regulation of blood glucose in the crayfish *Procambarus clarkii*. *In vitro*, both L-Enk and DA reduced the rate of release of CHH from isolated eyestalk neuroendocrine tissue. These results suggest that both L-Enk and DA act to inhibit the release of CHH by affecting the eyestalk neuroendocrine complex and that the

Table 2. Identification and localization of opioid-like peptides in insects

Peptide	Species	Localization	Techniques	References
Leu-Enk-like	*Calliphora vomitoria*	Brain, thoracic ganglia, retrocerebral complex	Immunocytochemistry	50
	Locusta migratoria	Brain	RIA	42
	Eristalis aeneus	Brain	Immunocytochemistry	48
	Manduca sexta	Brain	Immunocytochemistry	49
Met-Enk-like	*Calliphora vomitoria*	Brain, thoracic ganglia, retrocerebral complex	Immunocytochemistry	50
	Drosophila melanogaster	Brain, ovaries, testes	Immunocytochemistry	47, 57
	Leucophaea maderae	Corpora cardiaca	Immunocytochemistry	46
	Locusta migratoria	Brain, corpora cardiaca, spermatogonia, spermatids, spermatocytes, ovarian follicles	Immunocytochemistry, RIA	42, 52, 54, 55
	Periplaneta americana	Brain, nerves leading to corpora cardiaca and corpora allata, midgut endocrine cells	Immunocytochemistry	45, 53
	Nauphoeta cinera	Midgut endocrine cells	Immunocytochemistry	54
	Sarcophage bullata	Spermatogonia, spermatids, spermatocytes, ovarian follicles	Immunocytochemistry	55, 56
	Schistocera gregaria	Brain, thoracic ganglia, subesophageal ganglia, abdominal ganglia	Immunocytochemistry	51
Met-7	*Calliphora vomitoria*	Brain, thoracic ganglia, retrocerebral complex	Immunocytochemistry	50
Met-8	*Calliphora vomitoria*	Brain, thoracic ganglia, retrocerebral complex	Immunocytochemistry	50
Endorphin	*Bombyx mori*	Neurosecretory cells in the pars intercerebralis	Immunocytochemistry	59
	Calliphora vomitoria	Neurosecretory cells in the pars intercerebralis	Immunocytochemistry	60
	Eristalis aeneus	Neurosecretory cells in the pars intercerebralis	Immunocytochemistry	60
	Leptinotarsa decemlineata	Neurosecretory cells in the pars intercerebralis	Immunocytochemistry	58
α-Endorphin	*Bombyx*	Median cells of suboesophageal ganglion	Immunocytochemistry	61–63
	Thaumetopoea gityocampa	Median cells of suboesophageal ganglion	Immunocytochemistry	61, 62
MSH	*Leucophaea maderae*	Corpora cardiaca, corpora allata	Immunocytochemistry	65
	Locusta migratoria	Cell bodies and nerve fibres in the central nervous system	Immunocytochemistry	64
	Sarcophaga bullata	Cell bodies and nerve fibres in the central nervous system	Immunocytochemistry	64
ACTH	*Leucophaea maderae*	Corpora cardiaca, corpora allata, Malpighian tubules	Immunocytochemistry, ACTH antiserum	65, 66
	Locusta migratoria	Corpora cardiaca	Immunocytochemistry	64

RIA: radioimmunoassay.

enkephalinergic inhibitory neuron follows the dopaminergic inhibitory neuron in the chain of neurons which leads to the neuroendocrine cells that secrete CHH.

Reproduction. The potential involvement of an endogenous opioid system in the regulation of ovarian development in the *Uca pugilator* was investigated *in vivo*[41]. Injection of synthetic Met-Enk into female crabs slowed ovarian maturation significantly. The inhibition was dose-dependent. The mean ovarian index and the mean oocyte diameter of the crabs that received the opioid were significantly smaller than the corresponding values for the saline-injected concurrent control specimens. Injection of naloxone produced dose-related ovarian maturation[41].

Insecta

Localization and distribution

The first demonstration of the presence of an opioid in an invertebrate came from an RIA study with the brain of a locust, *Locusta migratoria*[42]. No β-endorphin was found but Leu-Enk and Met-Enk were both present, with Met-Enk present in larger quantity than Leu-Enk. Since then immunocytochemical studies based on antisera raised against opioid peptides have been carried out in about 35 species of insects belonging to eight orders[43,44]. Table 2 summarizes the results so far obtained regarding the localization and distribution of opioid peptides in insects as well as the methods employed to detect these peptides[45–66]. A few investigators have demonstrated the presence of opioid-binding sites in insect neural tissues. Pert and Taylor[68] showed that membranes prepared from heads of *Drosophila melanogaster* bind Leu-Enk, but naloxone does not bind well to these membranes. D-ala-2-met-5-enkephalinamide (DAMA) also binds extensively to cerebral ganglion membrane suspensions from *Leucophaea maderae*[67]. In the same insect, high-affinity stereospecific opioid binding was demonstrated in midgut preparations[69]. Santoro *et al.*[70] reported that tissues of *Drosophila* appear to contain kappa-like opioid receptors. These results suggest that insects possess both opioid substances and their receptors, thus revealing the existence of a complex opioid mechanism similar to mammals. Recently, Lamango and Isaac[71] characterized an endopeptidase from the heads of *Musca domestica* that can degrade a range of neuropeptides, including enkephalins and substance P.

Functions

Neuromodulation. Morphological evidence suggests the existence of specific enkephalinergic pathways within the central nervous system (CNS) of insects[43]. Immunoreactive nerve fibres originating in the brain were seen in nerves issuing from the corpora cardiaca (CC) and running along pathways leading to the anterior sympathetic nervous system and corpora allata (CA). These opioid peptides may be functioning as neurotransmitter or neuromodulatory substances[43]. In *Calliphora vomitoria*, the heart has a direct enkephalinergic innervation, particularly by Met-7 immunoreactive terminals in the vicinity of the ostia; the enkephalin-like material presumably acts as a neuromodulator of the intrinsic heart rhythm[43]. In addition, in this insect the phasic secretory activity of the CA is regulated by enkephalin-like peptide[43].

Neurohormone. The presence of enkephalins in the neurohaemal areas of the dorsal neural sheath of the thoracic ganglia of *Calliphora vomitoria* suggests a hormonal role for these peptides[43]. Furthermore, a close association between 5-HT and a variety of peptidergic substances, including Leu-Enk and Met-7, was observed. Met-7 at a concentration of as low as 10^{-12} M displayed a powerful stimulatory influence on salivary gland secretion in the rat. In contrast, Leu-Enk and Met-Enk were completely without effect. Naloxone failed to inhibit this response of the rat on salivary glands to Met-7. Thorpe and Duve[43] suggested that Met-7 may act through kappa receptors, where naloxone has only a weak binding capacity.

Behaviour. A behavioural role for an enkephalin in regulating locomotor activity in the cockroach *Leucophaea maderae* was demonstrated by Ford *et al.*[72]. When the enkephalin analogue D-ala-2-met-5-enkephalinamide (DAMA) was applied topically to the cerebral ganglia, it resulted in a decrease in locomotor activity. Naloxone can block this effect as well as the depressant effect of morphine on locomotor activity. In contrast, another analogue, D-ala-2-leu-5-enkephalinamide (DALA) and also dynorphin increased the locomotor activity. This acceleratory effect can also be antagonized by concomitant naloxone treatment. In another cockroach, *Periplanata americana*, enhanced locomotor activity was induced by morphine, a mu opioid agonist[73]. In the same species, it was shown that the kappa opiate agonist U-50,488H was able to stimulate the ingestive responses of free-feeding animals. These results on opioid-induced behaviours indicate the presence of multiple opiate receptor types in insects.

Thermoregulation. Survival of an animal is dependent on its ability to avoid thermal extremes. Many ectotherms show at least a primitive ability to control their body temperature. The behavioural thermoregulatory responses of at least some insects are sensitive to opiates. Kavaliers and Hirst[74] studied the thermal responses of larvae of the mealworm *Tenebrio molitor*. Fourth instar larvae were injected with morphine, Met-Enk, β-endorphin,

Table 3. Identification and localization of opioid-like peptides in molluscs

Peptide	Species	Localization	Techniques	References
Leu-Enk-like	*Aplysia californica*	Central nervous system	Immunocytochemistry	83
	Cepaea nemoralis	Brain neurons, central nervous system	Immunocytochemistry enzyme assay	84, 89
	Lepidopleurus asellus	Central nervous system	Immunocytochemistry	84
	Lymnaea magnalis	Central nervous system	Immunocytochemistry	83
	Lymnaea stagnalis	Central nervous system, peripheral nervous system	Immunocytochemistry	82, 87
	Mytilus edulis	Pedal ganglia, cerebral ganglia	Immunocytochemistry, HPLC, amino acid sequencing	90, 91
Met-Enk-like	*Achatina fulica*	Cerebral ganglia, visceral ganglia, nerves and ganglion cells of posterior intestine, parietal ganglia	Immunocytochemistry	80, 81
	Acmaea testudinalis	Peripheral nervous system	Immunocytochemistry	84
	Aplysia californica	Central nervous system	Immunocytochemistry	83
	Cepaea nemoralis	Brain neurons, peripheral nervous system	Immunocytochemistry	84
	Helix aspersa	Gonads, peripheral nervous system	Immunocytochemistry	84, 85
	Helix pomatia	Central nervous system	Immunocytochemistry	83
	Lepidopleurus asellus	Central nervous system	Immunocytochemistry	84
	Littorina littorea	Peripheral nervous system	Immunocytochemistry	84
	Lymnaea magnalis	Central nervous system	Immunocytochemistry	83
	Lymnaea stagnalis	Cerebral ganglia, pedal ganglia, central nervous system	Immunocytochemistry, HPLC	79, 87
	Mytilus edulis	Pedal ganglia, cerebral ganglia	Immunocytochemistry, HPLC, amino acid sequencing	90, 91
	Octopus vulgaris	Vena cava neurohaemal organ, neurosecretory cells of the brain	Immunocytochemistry	93, 94
Met-Enk-Arg^6-Phe^7	*Mytilus edulis*	Cerebral ganglia, pedal ganglia	Immunocytochemistry, HPLC	92
	Octopus vulgaris	Vena cava neurohaemal organ, neurosecretory cells of the brain	Immunocytochemistry	94, 95
ACTH	*Lymnaea stagnalis*	Central nervous system	Immunocytochemistry	87
1–24 ACTH	*Helix aspersa*	Retina, ganglionic cells, some spermatogenic cells	Immunocytochemistry	85, 88
	Planorbis corneus	Brain	Immunocytochemistry	86
17–39 ACTH	*Helix aspersa*	Young oocytes, nerve fibres surrounding each gonadal acinus and hermaphrodite duct	Immunocytochemistry	85
α-MSH	*Helix aspersa*	Retina, ganglionic cells, young oocytes	Immunocytochemistry	85, 88
	Lymnaea stagnalis	Central nervous system	Immunocytochemistry	87
	Planorbis corneus	Brain	Immunocytochemistry	86

HPLC: high-performance liquid chromatography.

or naloxone. Their behaviour and temperature preference were recorded following their return to pretreatment locations in a thermal gradient with a low temperature of 20°C and a high temperature of 40°C. Prior to drug treatment, the larvae showed a preference for a temperature of about 32°C. After injection, animals given morphine or Met-Enk moved towards higher temperatures whereas those receiving naloxone did not. Larvae receiving β-endorphin displayed no consistent changes in thermal preference. The changes in preferred

temperatures that took place were evident for approximately 1 h, after which animals returned to the pretreatment preference temperature[74].

Immune response. Among recent developments in the elucidation of opioid activities in nonnervous cells and tissues, endogenous opioids in the immune system are receiving the greatest attention in insects. The endogenous opioids in *Leucophaea maderae* appear to influence the autoregulatory activities of immunocompetent cells. These activities include the adherence and migratory behaviour of these cells[75]. There is evidence for the presence in immunocytes, as well as in cell-free haemolymph, of Met-Enk, which appears to be involved in immunoregulatory activities, perhaps by interaction with a special subtype of delta receptor. Moreover, the naloxone has an inhibitory effect on these processes, which indicates that they are receptor-mediated[75]. Thus, there appears to be a link between the neuroendocrine and the 'immune' system in insects[76].

Mollusca

Identification and localization

Tremblay *et al.*[77] seem to have been the first investigators to demonstrate the action of an opiate peptide in a mollusc. Injection of morphine into the gastropod *Aplysia californica* decreased the available neurotransmitter, acetylcholine. Erdélyi *et al.*[78] demonstrated that in the snail *Helix pomatia* Leu-Enk induces hyperpolarization of some neurons in the brain, and naloxone administration antagonizes the action of this opioid. These results suggest the presence of an endogenous opioid system in gastropod molluscs. By use of immunocytochemical techniques, a very detailed study of the immunoreactive material present in the central nervous system of the pond snail *Lymnaea stagnalis* has been performed by Schot *et al.*[79]. However, Met-Enk-like material was observed only in the cerebral and pedal ganglia. In each cerebral ganglion a cluster of 10–15 immunoreactive cells lies in the mediolateral part of the ganglion; another cluster (6–10 cells) is just dorsal to the lateral lobe. In addition, two positively reacting cells were observed in the medioventral area. Numerous positive axonal terminals were seen along the entire periphery of the cerebral commissure. In each pedal ganglion a group of 31–40 immunoreactive cells, located laterally, show a positive reaction for Met-Enk. A second group (8–10 cells) lies in the ventromedial area of the ganglion. Such immunocytochemical studies have been extended to other molluscs. Table 3 summarizes the evidence for the presence of opioids in various molluscs so far investigated[79–81,83–95]. Using HPLC, Leung *et al.*[82] identified a substance in the abdominal ganglion of *Aplysia californica* that was able to displace DAMA. In the nervous system of *Mytilus*, evidence was obtained for two classes of opioid-binding sites in nervous tissue[96,97]. It was suggested that the two binding sites may be associated with presynaptic and postsynaptic receptor sites for the opioids, or possibly with different pharmacological subtypes of receptors[98].

Functions

Behaviour. In the land snail *Helix pomatia* morphine induces a state of immobilization and muscle rigidity, resulting in a loss of the righting reflex. These effects are reversible by naloxone treatment[99,100]. This behavioural effect of morphine is reduced as treatments progress for four days and reoccurs if a higher dose of morphine is given on the fifth day, thus demonstrating that tolerance has occurred. It was suggested that in these molluscan behavioural responses enkephalins can mediate sensory neurotransmission[101]. Electrophysiologically, opioids are known to elicit a variety of actions in molluscan neurons, including both excitatory and inhibitory responses, some of which are naloxone-reversible and some that are not[102–104]. The first report of the effect of an opioid in an invertebrate appears to be that of Stefano and Hiripi[105], in which they describe an increase in the intraganglionic level of DA in response to Met-Enk and Leu-Enk in *Mytilus edulis*. Naloxone blocked the effects of Leu-Enk and Met-Enk. These responses displayed a time-dependent desensitization to Met-Enk. With the freshwater bivalve *Anodonta cygnea*, Stefano and Hiripi[106] also found an increase in the DA level in the cerebral ganglia after injection of Met-Enk or morphine. Furthermore, both agents increased the levels of cyclic GMP and depressed the levels of cyclic AMP. The pharmacological effects on DA and the cyclic nucleotide levels are blocked by prior treatment of the cerebral ganglia with naloxone. The data suggest that the decrease in the cyclic AMP concentration caused by the opiates may be a result of inhibition of adenylate cyclase[107]. Aiello and Hager[108] showed that the application of morphine to the visceral ganglion of *Mytilus edulis* is cilioexcitatory to the lateral cells of the gills, and this is most likely a result of decreasing the release of DA from cilioinhibitory fibres in the ganglion. Acetylcholine (ACh) has been proposed as the excitatory neurotransmitter for stimulation of tonic contraction of the anterior byssus retractor muscle of *Mytilus*. Morphine at concentrations of 10^{-9}–10^{-4} M increased markedly the contraction induced by ACh (ref. 109).

In the snail *Helix pomatia*, a number of identified neurons regulate visceral functions, and they react selectively to Leu-Enk and morphine; some become depolarized whereas others are hyperpolarized[110]. These peptides modulate the effect of ACh and 5-HT. It is known that a combination of low molecular weight neu-

rotransmitters and peptides modulates the elementary form of learning in *Helix*, such as regulation of various rhythmic processes, including movements of the pneumostoma[111]. In this snail, Leu-Enk modulates the 5-HT effect on the habituating central neurons through a cyclic 3′-5′-AMP system both *in situ* and *in vitro*. The habituating cells are connected to the regulation of various rhythmic processes, including pneumostoma (respiratory orifice) movements. In the pteropod mollusc *Clione limacina*, sensitivity to some tactile inputs can be decreased by naloxone, and it is suggested that in these behavioural responses enkephalins mediate sensory transmission[112].

Respiration. Moroz[113] reported that 5-HT activates the respiratory rhythm in the snail *Lymnaea stagnalis*, while enkephalin acts to limit this feedback system. In the marine gastropod *Aplysia californica*, which uses gill respiration, Met-Enk inhibits gill movements[114]. Furthermore, in this species it is known that 5-HT is able to activate gill movements[115]. Dyakonova[116,117] presented clear evidence for central systemic effects of Met-Enk and 5-HT in *Helix pomatia*. She showed that Met-Enk and 5-HT are able to induce slow oscillation of the membrane potential in a group of respiratory neurons. Synchronization of the activity between the neurons in the cluster appeared to be due to an increase of electrical coupling between the cells[117]. This suggested that Met-Enk might be involved in modulation of respiratory programmes, including the central respiratory generator. For another snail, *Lymnaea stagnalis*, Moroz and Winlow[118] reported that Met-Enk inhibits the 5-HT-induced respiratory rhythm during the first few minutes after this opioid is administered. Later, Met-Enk (5–30 min after its administration) induces slow oscillation of the membrane potential in central neurons related to the respiratory programme. More recently, Dyakonova *et al.*[119] studied the effect of Met-Enk on electrical coupling between molluscan neurons, using isolated brains of *Helix* and *Lymnaea*. In the presence of both 5-HT and Met-Enk, nonrectifying electrical coupling is strongly facilitated between identified respiratory neurons in *Helix*, while coupling between 5-HT-containing ciliomotoneurons in *Lymnaea* is facilitated by Met-Enk alone. These data suggest that Met-Enk can modulate different groups of electrically coupled cells and may be involved in coordination of motor patterns[119].

Feeding. Administration of morphine resulted in significant dose-dependent increases in the ingestive responses of food-deprived slugs *Limax maximus*, and in the initiation of feeding in satiated animals[120]. These effects could be blocked by naloxone, with naloxone by itself causing a significant decrease in feeding by food-deprived slugs. These results suggest that opioids are involved in the control of feeding behaviour of the slugs.

Thermoregulation. Terrestrial snails are ectotherms, their body temperatures being dependent on the ambient temperature. The terrestrial snail *Cepaea nemoralis*, when placed on a 40°C hot plate, lifts the anterior portion of its foot. The latency of this response is affected by morphine and naloxone[121]. Administration of morphine resulted in a significant dose-dependent increase in the temperature selected by the snail in a thermal gradient. These thermoregulatory effects could be blocked and reversed by naloxone, with the opiate antagonist by itself causing a significant decrease in the preferred temperature[122]. After 6–10 days of daily administration of morphine, *Cepaea nemoralis* displayed tolerance; the effect of morphine decreased and the thermal preferences of the snails administered morphine became similar to those of saline-treated individuals. These results indicate that the opiate system plays a thermoregulatory role in molluscs. With the same species of snail, Kavaliers and Hirst[74] observed temperature responses similar to those produced by morphine when the snails were treated with Met-Enk or β-endorphin. Naloxone antagonized the effects of these peptides. Recently, Kavaliers and Ossenkopp[123] reported that repeated naloxone treatments and exposure to weak (60 Hz) magnetic fields brought about a hypoalgesic response to thermal nociceptive stimuli in *Cepaea nemoralis*.

Opioids in the aging process. Age-related changes in the opioid system of *Mytilus edulis* were detected by Stefano[124]. He studied the opioid binding profile in aging pedal ganglia. The high-affinity binding site density is significantly lower in older animals than in younger ones, whereas the low-affinity site density remains unchanged. The estimated Met-Enk and Met-Enk-Arg^6-Phe^7 levels were significantly higher for the older animals than for the younger ones. The decrease in the high-affinity binding site density and the corresponding increase in endogenous enkephalin levels suggest the existence of an opioid compensatory mechanism associated with the aging process[125].

Immune responses. Ottaviani *et al.*[126] have demonstrated that in the gastropod *Planorbarius corneus* ACTH (1–24) and β-endorphin can stimulate haemocyte locomotion and phagocytosis under stress.

Annelida

Localization and distribution

The earliest information concerning the presence of opioid peptides in annelids was produced by Alumets *et al.*[127] and Rémy and Dubois[128], who studied earthworms. Using immunocytochemistry, Alumets *et al.*[127] demonstrated the presence of immunoreactive Leu-Enk-like material in the cerebral ganglia and cerebral connectives

Table 4. Identification and localization of opioid-like peptides in annelids

Peptide	Species	Localization	Techniques	References
Leu-Enk-like	*Eisenia foetida*	Neurons of the brain	Immunocytochemistry	132
	Haemopis marmorata	Segmental ganglia	Immunocytochemistry	135, 137
	Lumbricus terrestris	Cerebral ganglia, cerebral connectives, suboesophageal ganglia, ganglia of the nerve cord, central nervous system, gut, seminal vesicle, body wall tissue, epidermis, nerve cord	Immunocytochemistry, RIA	127, 130, 133
	Nereis diversicolor	Central nervous system, peripheral nerves, oocytes, spermatocytes	Immunocytochemistry	140–144
	Theromyzon tessulatum	Central nervous system, proboscis, foregut	Immunocytochemistry	138
Met-Enk-like	*Eisenia foetida*	Neurons of the brain	Immunocytochemistry	132
	Hirudo medicinalis	Brain neurosecretory cells, fibres within the neuropil of the brain	Immunocytochemistry	136
	Lumbricus terrestris	Central nervous system, gut, seminal vesicles, body wall tissue, epidermis, nerve cord	Immunocytochemistry, RIA	130, 133
	Nereis diversicolor	Central nervous system, peripheral nerves	Immunocytochemistry	140–142
α-Endorphin	*Dendrobaena subrubicunda*	Suboesophageal ganglia	Immunocytochemistry	128
	Nereis diversicolor	Central nervous system, peripheral nerves	Immunocytochemistry	140–142
β-Endorphin	*Lumbricus terrestris*	Cerebral ganglia, cerebral connectives, suboesophageal ganglia, ganglia of the nerve cord	Immunocytochemistry	127, 131
	Nereis diversicolor	Central nervous system, peripheral nerves	Immunocytochemistry	140–142
α-Neoendorphin	*Theromyzon tessulatum*	Central nervous system, proboscis	Immunocytochemistry	138
Dynorphin	*Eisenia foetida*	Neurons of the brain	Immunocytochemistry	132
	Nereis diversicolor	Central nervous system, peripheral nerves, oocytes, spermatocytes	Immunocytochemistry	140–142
	Theromyzon tessulatum	Central nervous system, proboscis	Immunocytochemistry	138
ACTH-like	*Dendrobaena subrubicunda*	Cerebral ganglion	Immunocytochemistry	129
	Lumbricus terrestris	Epidermis, nerve cord	Immunocytochemistry	134
	Nereis diversicolor	Central nervous system	Immunocytochemistry	144
	Theromyzon tessulatum	Brain	Immunocytochemistry	139
MSH	*Lumbricus terrestris*	Epidermis, nerve cord	Immunocytochemistry	134
α-MSH	*Theromyzon tessulatum*	Brain	Immunocytochemistry	139
β-MSH	*Nereis diversicolor*	Central nervous system	Immunocytochemistry	144

RIA: radioimmunoassay.

of *Lumbricus terrestris*. They also observed in some, but not all, of the specimens immunoreactive Leu-Enk-like material in the suboesophageal ganglion and the ganglia of the nerve cord. Cells immunoreactive to β-endorphin have a distribution similar to those of the enkephalin, but differ in number and size. A summary of such localization studies[127–144] as well as distribution of opioid peptides in other annelids are presented in Table 4.

Physiological evidence for the existence of opioid receptors in the leech has been reported by Flanagan and

Zipser[137]. Homogenates of the nerve cord of the leech, *Haemopis marmorata*, display saturable binding of 3H-(D-Ala^2-)-Met-Enk.

Functions

Behaviour. Leeches respond behaviourally to externally applied morphine and to morphine injections. Kaiser[145] reported that morphine produces a biphasic hyperkinetic–quiescent locomotory response in free-swimming *Hirudo medicinalis*. Flanagan and Zipser[137] found a similar biphasic response in another leech, *Haemopis marmorata*, that was injected with morphine. Control animals injected with saline were initially hyperactive after injection, but rapidly reattached themselves to the wall of the holding tank and remained highly responsive to subsequent tactile stimulation. Morphine-injected leeches were also initially hyperactive, but after 5–10 min they fell to the bottom of the tank and remained unresponsive to all but the most extreme tactile stimulation for a period of 4 days. These animals fully recovered by day 5. This behavioural effect appears to have been due to inhibition of synaptic transmission by morphine. Gardner and Walker[146] reported an inhibitory effect of morphine on the Retzius cells of *Hirudo medicinalis*. Other evidence from *Hirudo medicinalis* and *Haemopis marmorata* suggests that the enkephalins are acting either as primary neuromodulators or as sensory neurotransmitters[147].

Feeding. Kaloustian and Rzasa[131], using radioimmunoassay found the highest concentrations of Met-Enk-like, Leu-Enk-like and β-endorphin-like material in the region of the gut which contains high digestive enzyme activity of *Lumbricus terrestris*. Isolated gut treated with these opioids showed significant decreases in its contractile properties. It appears that these opioids exert their effects by modulating the action of DA, resulting thereby in an increased transit time for food in areas of the gut with high digestive enzyme activities, hence enhancing the digestive process.

Reproduction. Experiments were done with *Nereis diversicolor* to determine whether dynorphin may have a role in the control of reproduction[144]. The basis for these experiments was the presence of this opioid in the infracerebral neurohemal area of the cerebral ganglia. Parapodia in which spermatogenesis was proceeding were incubated in sea water containing dynorphin. However, no effect on spermatogenesis was observed. Thus, it seems likely that if dynorphin-like peptides have a role in reproduction, they do not act directly on the spermatocytes, in contrast to the direct action of 'nereidin' (a brain neuropeptide that inhibits the maturation of spermatocytes in nereids)[148]. Therefore, if these dynorphin-like peptides have any role in reproduction, it is only to function as neurotransmitters to control the release of reproductive hormones.

Other invertebrates

Immunocytochemical evidence for the occurrence of opioid peptides in invertebrates other than annelids, arthropods and molluscs is rather scarce. Compounds similar to ACTH (1–39) and β-endorphin have been detected in the protozoan *Tetrahymena pyriformis*[149]. Roth *et al.*[150] reported that opioid peptides alter the feeding behaviour of the amoeba, with the response being blocked by naloxone. Venturini *et al.*[151,152] detected Met-Enk-like immunoreactivity in neuronal perikarya and in neuropil in the planarian *Dugesia gonocephala*. In this planarian naloxone induces a dose-dependent rise in the cAMP level while morphine decreases cAMP. Morphine treatment decreases the motility of this planarian whereas naloxone treatment induces screw-like hyperkinesis. Treatment with dopaminergic agonists (L-dopa, apomorphine) induces typical screw-like hyperkinesis, whereas after treatment with reserpine and haloperidol all the planarians became motionless, a response similar to that after treatment with morphine. A model for interactions between dopaminergic and enkephalinergic neurotransmission and/or neuromodulation was proposed that involved inhibitory modulation of opiates on DA release[152]. Romero *et al.*[153] reported that Met-Enk decreased the mitotic rate of adult planarians during regeneration while substance P was stimulatory. Immunocytochemical studies showed the presence of Met-Enk within the central and peripheral nervous systems of this planarian.

Piccoli *et al.*[13] made a study of the presence of endogenous Met-Enk and Leu-Enk in 11 species of marine organisms ranging from sponges to tunicates. Extracts from whole-animal tissues of the sponge *Dysidea avaria*, the cnidarians *Anemonia sulcata*, *Pelagia noctiluca*, *Rhyzostoma pulmo*, and the tunicate *Ciona intestinalis* were found to contain opioid peptides. Rémy and Brossard[154] observed immunoreactivity to Met-Enk in the cephalic nervous system of two species of nemerteans, *Lineus viridis* and *Lineus ruber*, but in the ventral ganglia, Met-Enk-like material is rather scarce. Colon-Urban *et al.*[155] suggested that endogenous opioids in various bryozoans (*Bugula neritina* and *Membranipora* sp.) play a role in the organisms' motor mechanism. Morphine inhibits tentacle activity, an effect which can be blocked by prior or concomitant naloxone application. Engbretson and Chamberlain[156] reported positive immunoreactivity to Leu-Enk in fibres of the optic medulla of the horseshoe crab *Limulus polyphemus*.

Conclusions

Although the study of endogenous opioid peptides in invertebrates is still at an early stage, the data presented in this review support the view that such molecules are present throughout the invertebrate phyla and have a

long evolutionary history. The application of immunocytochemical techniques has greatly influenced opioid research in invertebrates. Despite an impressive amount of research in invertebrates, for many of the species studied the functional significance of these opioid peptides and the mechanisms underlying their action remain elusive. The biochemical, physiological and pharmacological characterization of opioid peptides needs a careful study in order to strengthen our knowledge about their actual role in the physiology of invertebrates. Another important area that merits much attention is the identification and purification of the receptors for opioid peptides in invertebrates. Elucidation of the molecular structures of the opioid receptors will open new avenues for the evaluation of the molecular aspects of opioid receptor function.

A challenging phenomenon of current opioid research in higher animals, including mammals, is the colocalization of opioid peptides with nonopioids and classical neurotransmitters in the same nerve terminal. The rapidly increasing data on this issue show a bewildering complexity, which is presently difficult to understand in the mammalian nervous system. Simple models using invertebrates may be more suitable to unravel the nature of the putative interactions between opioid peptides and classical neurotransmitters. Finally, invertebrate opioid research may have practical applications to help control harmful organisms, such as insect pests and barnacles. Knowledge of the neurochemistry of the opioids in the invertebrates could lead to the discovery of novel ways to inhibit reproduction and growth of these organisms. If opioid peptides are important regulatory agents in metabolism, then it might be possible to devise methods of breaking particular links in the synthesis of these compounds, and research directed along these lines could provide us with much-needed compounds for the control of invertebrate pests.

1. Frederickson, R. C. A. and Geary, L. E., *Prog. Neurobiol.*, 1982, **19**, 19–69.
2. Dixon, J. E., Haun, R. S., Minth, C. D. and Nichols, R., in *Neuropeptides and their Peptidases* (ed. Turner, A. J.), Weinheim, New York, 1987, pp. 10–14.
3. Hughes, J., Smith, T. W., Kosterlitz, H. W., Fothergill, L. A., Morgan, B. A. and Morris, H. R., *Nature*, 1975, **258**, 577–579.
4. Kosterlitz, H. W., *Proc. R. Soc. London*, 1985, **B225**, 27–40.
5. Olson, G. A., R. D., Kastin, A. J. and Coy, D. H., *Peptides*, 1981, **2**, 349–369.
6. Olson, G. A., Olson, R. D., Kastin, A. J. and Coy, D. H., *Peptides*, 1982, **3**, 1039–1072.
7. Pert, C. B., Aphoshian, D. and Snyder, S. H., *Brain Res.*, 1974, **75**, 356–361.
8. Simantov, R., Goodman, R., Aphoshian, D. and Snyder, S. H., *Brain Res.*, 1976, **111**, 204–211.
9. Mancillas, J. R., McGinty, J. F., Selverston, A. I., Karten, H. and Bloom, F. E., *Nature*, 1981, **293**, 576–578.
10. Jaros, P. P., Dircksen, H. and Keller, R., *Cell Tissue Res.*, 1985, **241**, 111–117.
11. Fingerman, M., Hanumante, M. M., Kulkarni, G. K., Ikeda, R. and Vacca, L. L., *Cell Tissue Res.*, 1985, **241**, 473–477.
12. Marino, G., Palmisano, A., DiMarzo, V. and Melck, D., *Peptides*, 1985, **6** (Suppl. 3), 403–406.
13. Piccoli, R., Melck, D., Spagnuolo, A., Vescia, S. and Zanetti, L., *Comp. Biochem. Physiol.*, 1985, **C80**, 237–240.
14. Dircksen, H., Stangier, J. and Keller, R., *Gen. Comp. Endocrinol.*, 1987, **66**, 42.
15. Dircksen, H., in *Frontiers in Crustacean Neurobiology, Advances in Life Sciences* (eds Wiese, K., Krenz, W.-D., Tautz, J., Reichert, H. and Mulloney, B.), Birkhäuser, Basel, 1990, pp. 485–491.
16. Amiard-Triquet, C., Amiard, J.-C., Ferrand, R., Andersen, A. C. and Dubois, M. P., *Ecotoxicol. Environ. Safety*, 1986, **11**, 198–209.
17. Lüschen, W., Buck, F., Willig, A. and Jaros, P. P., *Proc. Natl. Acad. Sci. USA*, 1991, **88**, 8671–8675.
18. Rothe, H., Lüschen, W., Asken, A., Willig, A. and Jaros, P. P., *Comp. Biochem. Physiol.*, 1991, **C99**, 57–62.
19. Leung, M. K., Kessler, H., Whitefield, K., Murray, M., Martinez, E. A. and Stefano, G. B., *Cell. Mol. Neurobiol.*, 1987, **7**, 91–96.
20. Arbiser, Z. K. and Beltz, B. S., *J. Comp. Neurol.*, 1991, **306**, 417–424.
21. Fingerman, M., Sarojini, R., Mary, A., Mary, V. and Nagabhushanam, R., *Am. Zool.*, 1994, **34**, 22A.
22. Hara, M., Kouyama, H., Watabe, S. and Yago, N., *Zool. Sci. (Tokyo)*, 1985, **2**, 971.
23. Coletti-Previero, M.-A., Mattras, H., Zwilling, R. and Previero, A., *Neuropeptides*, 1985, **6**, 405–415.
24. Watabe, S., Hara, M. and Yago, N., *Zool. Sci. (Tokyo)*, 1985, **2**, 971.
25. Maldonado, H. and Miralto, A., *J. Comp. Physiol.*, 1982, **147**, 455–459.
26. Martinez, E. A., Murray, M., Leung, M. K. and Stefano, G. B., *Comp. Biochem. Physiol.*, 1988, **C90**, 89–93.
27. Lozada, M., Romano, A. and Maldonado, H., *Pharmacol. Biochem. Behav.*, 1988, **30**, 635–640.
28. Valeggia, C., Fernandez-Duque, E. and Maldonado, H., *Brain Res.*, 1989, **481**, 304–308.
29. Brunner, D. and Maldonado, H., *J. Comp. Physiol.*, 1988, **A162**, 687–694.
30. Godoy, A. M. and Maldonado, H., *Pharmacol. Biochem. Behav.*, 1995, **50**, 445–451.
31. Fingerman, M., in *Endocrinology of Selected Invertebrate Types* (eds. Laufer, H. and Downer, R. G. H.), A. R. Liss, New York, 1988, pp. 357–374.
32. Quackenbush, L. S. and Fingerman, M., *Comp. Biochem. Physiol.*, 1984, **C79**, 77–84.
33. Butler, T. A. and Fingerman, M., *Am. Zool.*, 1985, **25**, 102A.
34. Martinez, E. A., Vassell, D. and Stefano, G. B., *Comp. Biochem. Physiol.*, 1986, **C83**, 77–82.
35. Kulkarni, G. K. and Fingerman, M., *Comp. Biochem. Physiol.*, 1986, **C84**, 219–224.
36. Kulkarni, G. K. and Fingerman, M., *Comp. Biochem. Physiol.*, 1991, **C99**, 47–52.
37. Kulkarni, G. K. and Fingerman, M., *Pigment Cell Res.*, 1987, **1**, 51–56.
38. Fingerman, M. and Nagabhushanam, R., *Comp. Biochem. Physiol.*, 1992, **C102**, 343–352.
39. Kallen, J. L., Rigiani, N. R. and Trompenaars, H. J. A. J., *Biol. Bull.*, 1988, **175**, 137–143.
40. Sarojini, R., Nagabhushanam, R. and Fingerman, M., *Gen. Comp. Endocrinol.*, 1995, **97**, 160–170.
41. Fingerman, M., Sarojini, R. and Nagabhushanam, R., *Neuroendocrinology*, 1994, **60** (Suppl. 1), 74.
42. Gros, C., Lafon-Cazal, M. and Dray, F., *C. R. Acad. Paris*, 1978, **D287**, 647–650.
43. Thorpe, A. and Duve, H., in *Progress in Comparative Endocrinology* (eds Epple, A., Scanes, C. G. and Stetson, M. H.), Wiley-Liss, New York, 1990, pp. 293–299.

44. Raabe, M., in *Recent Developments in Insect Neurohormones*, Plenum Press, New York, 1989, pp. 1–503.
45. Verhaert, P. and DeLoof, A., *J. Comp. Neurol.*, 1985, **239**, 54–61.
46. Hansen, B. L., Hansen, G. N. and Scharrer, B., *Cell Tissue Res.*, 1982, **225**, 319–329.
47. Pagés, M., Jiménez, F., Ferrus, A., Peralta, E., Ramirez, G. and Gelpi, E., *Neuropeptides*, 1983, **4**, 87–98.
48. El-Sahly, M., Abou-El-Ela, R., Falkner, S., Grimelius, L. and Wilander, E., *Regul. Peptides*, 1980, **1**, 187–204.
49. El-Sahly, M., Falkner, S., Kramer, K. J. and Speirs, R. D., *Cell Tissue Res.*, 1983, **232**, 295–317.
50. Duve, H. and Thrope, A., *Cell Tissue Res.*, 1988, **251**, 399–415.
51. Davenport, A. P. and Evans, P. D., *Brain Res.*, 1986, **383**, 319–322.
52. Romeuf, M. and Rémy, C., *Cell Tissue Res.*, 1984, **236**, 289–292.
53. Nishütsutsuji-Uwo, J. and Endo, Y., *Biomed. Res.*, 1981, **2**, 30–44.
54. Sehnal, F. and Zitnan, D., in *Progress in Comparative Endocrinology* (eds Epple, A., Scanes, C. G. and Stetson, M. H.), Wiley-Liss, New York, 1990, pp. 510–515.
55. Schoofs, L., Schrooten, S., Huybrechts, R. and DeLoof, A., *Gen. Comp. Endocrinol.*, 1988, **69**, 1–12.
56. DeLoof, A., in *Endocrinological Frontiers in Physiological Insect Ecology* (eds Sehnal, F., Zabza, A. and Denlinger, D. L.), Wroclaw Technical University Press, Wiroclaw, 1988, pp. 813–831.
57. Callaerts, P., Huybrechts, R. and DeLoof, A., *Int. J. Invert. Reprod. Develop.*, 1988, **13**, 147–156.
58. Veenstra, J. A., Romberg-Privee, H. and Schooneveld, H., *Cell Tissue Res.*, 1985, **240**, 535–540.
59. Rémy, C., *J. Physiol. (Paris)*, 1982, **18**, 514–522.
60. Duve, H. and Thorpe, A., *Cell Tissue. Res.*, 1983, **233**, 415–426.
61. Rémy, C., Girardie, J. and Dubois, M. P., *C. R. Acad. Sci. Paris*, 1978, **D286**, 651–653.
62. Rémy, C., Girardie, J. and Dubois, M. P., *Gen. Comp. Endocrinol.*, 1979, **37**, 93–100.
63. Takeda, S., Vieillemaringe, J., Geffard, M. and Rémy, C., *Cell Tissue Res.*, 1986, **243**, 125–128.
64. Schoofs, L., Jégou, S., Vaudry, H., Verhaet, P. and DeLoof, A., *Cell Tissue Res.*, 1987, **248**, 25–31.
65. Hansen, B. L., Hanse, G. N. and Scharrer, B., in *CRC Handbook of Comparative Aspects of Opioid and Related Neuropeptide Mechanisms* (ed. Stefano, G. B.), CRC Press, Boca Raton, Florida, 1986, vol. I, pp. 213–222.
66. Rafaeli, A., Moshitsky, P. and Applebaum, S. W., *Gen. Comp. Endocrinol.*, 1987, **67**, 1–6.
67. Stefano, G. B. and Scharrer, B., *Brain Res.*, 1981, **225**, 107–114.
68. Pert, C. D. and Taylor, D., in *Endogenous and Exogenous Opiate Agonists and Antagonists* (ed. Way, E. L.), Pergamon Press, New York, 1980, pp. 87–94.
69. Stefano, G. B., Scharrer, B. and Assanah, P., *Brain Res.*, 1982, **253**, 205–212.
70. Santoro, C., Hall, L. H. and Zukin, R. S., International Narcotic Research Conference, Falmouth, MA, 1985, Abst. 0–22.
71. Lamango, N. S. and Isaac, R. E., *Insect Biochem. Mol. Biol.*, 1993, **23**, 801–808.
72. Ford, R., Jackson, D. M., Tetrault, L., Torres, J. C., Assanah, P., Harper, J., Leung, M. K. and Stefano, G. B., *Comp. Biochem. Physiol.*, 1986, **C85**, 61–66.
73. Kavaliers, M., Guglick, M. A. and Hirst, M., *Life Sci.*, 1987, **40**, 665–672.
74. Kavaliers, M. and Hirst, M., in *CRC Handbook of Comparative Aspects of Opioids and Related Neuropeptide Mechanisms* (ed. Stefano, G. B.), CRC Press, Boca Raton, Florida, 1986, vol. I, pp. 139–152.
75. Scharrer, B., in *Advances in Neuroimmunology* (eds Stefano, G. B. and Smith, E. M.), Manchester University Press, Manchester, 1991, vol. I, pp. 1–5.
76. Nässel, R., *Cell Tissue Res.*, 1993, **273**, 1–29.
77. Tremblay, J. P., Schlapfer, W. T., Woodson, P. B. J. and Barondes, S. H., *Brain Res.*, 1974, **81**, 107–118.
78. Erdélyi, L., Joó, F. and Soós, J., in *Neuropeptides and Neural Transmission* (eds Marsan, C. A. and Traczyk, W. Z.), Raven Press, New York, 1980, pp. 271–280.
79. Schot, L. P. C., Boer, H. H., Swaab, D. F. and Noorden, S. V., *Cell Tissue Res.*, 1981, **216**, 273–291.
80. Noorden, S. V., Fritsch, H. A. R., Grillo, T. A. I., Polak, J. M. and Pearse, A. G. E., *Gen. Comp. Endocrinol.*, 1980, **40**, 375–376.
81. Takayanagi, H. and Takeda, N., *Comp. Biochem. Physiol.*, 1988, **A91**, 613–620.
82. Leung, M. K., Rozsa, S. K., Hall, A., Kuruvilla, S., Stefano, G. B. and Carpenter, D. O., *Life Sci.*, 1986, **38**, 1529–1534.
83. Elekes, K., Stefano, G. B. and Carpenter, D. O., *Cell Tissue Res.*, 1993, **272**, 329–341.
84. Sakharov, D., Nezlin, L., Moroz, L. and Elofsson, R., *Brain Res.*, 1993, **620**, 114–121.
85. Marchand, C. R. and Dubois, M. P., *Cell Tissue Res.*, 1986, **245**, 337–341.
86. Lichtensteiger, W., Felix, D. and Celio, M., in *Neurotransmitters in Invertebtates* (ed. Rozsa, K. S.), Akadémiai Kiadó, Budapest, 1980, pp. 459–475.
87. Leung, M. K., Boer, H. H., Minnen, J. V., Lundy, J. and Stefano, G. B., *Brain Res.*, 1990, **531**, 66–71.
88. Magdelaine, S. and Marchand, C. R., *Comp. Biochem. Physiol.*, 1993, **C105**, 255–261.
89. Saksida, L. M., Galea, L. A. M. and Kavaliers, M., *Peptides*, 1993, **14**, 763–765.
90. Stefano, G. B. and Martin, R., *Cell Tissue Res.*, 1983, **230**, 147–153.
91. Leung, M. K. and Stefano, G. B., *Proc. Natl. Acad. Sci. USA*, 1984, **81**, 955–958.
92. Stefano, G. B. and Leung, M. K., *Brain Res.*, 1984, **298**, 362–365.
93. Martin, R., Frösch, D., Weber, E. and Voigt, K. H., *Neurosci. Lett.*, 1979, **15**, 253–257.
94. Martin, R., Frösch, D., Weber, E. and Voigt, K. H., *Neuropeptides*, 1981, **2**, 141–150.
95. Voigt, K.-H. and Martin, R., in *CRC Handbook of Comparative Opioid and Related Neuropeptide Mechanisms* (ed. Stefano, G. B.), CRC Press, Boca Raton, Florida, 1986, vol. I, pp. 127–138.
96. Kream, R. M., Zukin, R. S. and Stefano, G. B., *J. Biol. Chem.*, 1980, **225**, 9218–9224.
97. Chapman, A., Gonzales, G., Burrowes, W. R., Assanah, P., Iannone, B., Leung, M. K. and Stefano, G. B., *Cell. Mol. Neurobiol.*, 1984, **4**, 143–155.
98. Stefano, G. B., Zukin, R. S. and Kream, R. M., *J. Pharmacol. Exp. Theor.*, 1982, **222**, 759–764.
99. Stefano, G. B., Vadaz, I. and Hiripi, L., *Experientia*, 1980, **36**, 666–667.
100. Burrows, W. R., Assanah, P. and Stefano, G. B., *Life Sci.*, 1983, **33**, 381–384.
101. Stefano, G. B., Rozsa, K. S. and Hiripi, L., *Comp. Biochem. Physiol.*, 1980, **C66**, 193–198.
102. Rózsa, K. S., *Comp. Biochem. Physiol.*, 1982, **C72**, 375–386.
103. Rózsa, K. S. and Solntseva, E. J., *Acta Physiol. Hung*, 1986, **67**, 429–433.
104. Lukowiak, K., Goldberg, J., Colmers, W. F. and Edstrom, J., in *CRC Handbook of Comparative Opioid and Related Neuropeptide Mechanisms* (ed. Stefano, G. B.), CRC Press, Boca Raton, Florida, 1986, vol. II, pp. 129–144.
105. Stefano, G. B. and Hiripi, L., *Life Sci.*, 1979, **24**, 1617–1622.
106. Stefano, G. B. and Hiripi, L., *Life Sci.*, 1979, **25**, 291–298.
107. Hiripi, L., in *CRC Handbook of Comparative Opioid and Related Neuropeptide Mechanisms* (ed. Stefano, G. B.), CRC Press, Boca Raton, Florida, 1986, vol. I, pp. 255–262.
108. Aiello, E. and Hager, E., in *CRC Handbook of Comparative Opioid and Related Neuropeptide Mechanisms* (ed. Stefano,

G. B.), CRC Press, Boca Raton, Florida, 1986, vol. II, pp. 233–242.
109. Bianchi, C. P. and Wang, Z., in *CRC Handbook of Comparative Opioid and Related Neuropeptide Mechanisms* (ed. Stefano, G. B.), CRC Press, Boca Raton, Florida, 1986, vol. II, pp. 59–64.
110. Salánki, J. and Rósza, K. S., in *CRC Handbook of Comparative Opioid and Related Neuropeptide Mechanisms* (ed. Stefano, G. B.), CRC Press, Boca Raton, Florida, 1986, vol. II, pp. 103–106.
111. Rózsa, K. S. and Dyakonova, T. L., *Comp. Biochem. Physiol.*, 1989, **C92**, 361–370.
112. Norekian, T. and Sakharov, D. A., *Sensory Systems*, 1991, **5**, 5–12.
113. Moroz, L. L., in *Signal Molecules and Behavior* (eds Winlow, W., Vinogradova, O. V. and Sakharov, D. A.), Manchester University, Manchester, 1991, pp. 101–123.
114. Lukowiak, K., Thornhill, J. A. and Endstrom, J., *Regul. Pept.*, 1982, **3**, ,303–312.
115. Weiss, S., Goldberg, J. I., Lukowiak, K. and Drummond, G. I., *J. Comp. Physiol.*, 1985, **B156**, 57–65.
116. Dyakonova, T. L., *Dokl. Acad. Nauk SSSR*, 1989, **308**, 1264–1269.
117. Dyakonova, T. L., in *Signal Molecules and Behavior* (eds Winlow, W., Vinogradova, O. V. and Sakharov, D. A.), Manchester University Press, Manchester, 1991, pp. 101–123.
118. Moroz, L. L. and Winlow, W., *Comp. Biochem. Physiol.*, 1993, **C106**, 179–187.
119. Dyakonova, T. L., Moroz, L. L. and Winlow, W., *Comp. Biochem. Physiol.*, 1993, **C106**, 93–101.
120. Kavaliers, M., Hirst, M. and Teskey, G. C., *Brain Res. Bull.*, 1985, **14**, 681–685.
121. Kavaliers, M., Hirst, M. and Teskey, G. C., *Science*, 1983, **220**, 99–100.
122. Kavaliers, M. and Hirst, M., *Neuropharmacology*, 1984, **23**, 1285–1289.
123. Kavaliers, M. and Ossenkopp, K. P., *Brain Res.*, 1993, **620**, 159–162.
124. Stefano, G. B., *Comp. Biochem. Physiol.*, 1982, **C72**, 349–352.
125. Stefano, G. B. and Leung, M. K., in *CRC Handbook of Comparative Opioid and Related Neuropeptide Mechanisms* (ed. Stefano, G. B.), CRC Press, Boca Raton, Florida, 1986, vol. II, pp. 199–210.
126. Ottaviani, E., Carselgrandi, E., Petraglia, F. and Franceschi, C., *Gen. Comp. Endocrinol.*, 1992, **87**, 334–360.
127. Alumets, J., Hakanson, R., Sundler, F. and Thorell, J., *Nature*, 1979, **279**, 805–806.
128. Rémy, C. and Dubois, M. P., *Experientia*, 1979, **35**, 137–138.
129. Aros, B., Wenger Vigh, B. and Vigh-Teichmann, I., *Acta Histochem.*, 1980, **66**, 262–268.
130. Rzasa, P. J., Kaloustian, K. V. and Prokop, E. K., *Comp. Biochem. Physiol.*, 1984, **C77**, 345–350.
131. Kaloustian, K. V. and Rzasa, P. J., in *CRC Handbook of Comparative Opioid and Related Neuropeptide Mechanisms* (ed. Stefano, G. B.), CRC Press, Boca Raton, Florida, 1986, vol. I, pp. 73–88.
132. Lkhider, M., Marcel, R. and Tramu, G., *Gen. Comp. Endocrinol.*, 1987, **65**, 457–468.
133. Gesser, B. P. and Larsson, L. I., *Cell Tissue Res.*, 1986, **246**, 33–37.
134. Al-Yousuf, S. A., in *Progress in Comparative Endocrinology* (eds Epple, A., Scanes, C. G. and Stetsor, M. H.), Wiley-Liss, New York, 1990, pp. 232–241.
135. Zipser, B., *Nature*, 1980, **283**, 857–858.
136. Osborne, N. N., Patel, S. and Dockray, G., *Histochemistry*, 1982, **75**, 573–583.
137. Flanagan, T. and Zipser, B., in *CRC Handbook of Comparative Opioid and Related Neuropeptide Mechanisms* (ed. Stefano, G. B.), CRC Press, Boca Raton, Florida, 1986, vol. I, pp. 165–180.
138. Verger-Bocquet, M., Malecha, J. and Tramu, G., *Comp. Endocrinol.*, 1987, **6**, 195–199.
139. Verger-Bocquet, M., Malecha, J. and Tramu, G., *Comp. Endocrinol.*, 1988, **7**, 21–27.
140. Dhainaut-Courtois, N., Dubois, M.-P., Tramu, G. and Masson, M., *Cell Tissue Res.*, 1985, **242**, 97–108.
141. Porchet, M., Dhainaut-Courtois, N., Cardon, C. and Bataille, M., in *Neurosecretion and the Biology of Neuropeptides* (eds Kobayashi, H., Bern, H. A. and Virano, A.), Springer, Berlin, 1985.
142. Dhainaut-Courtois, N., Tramu, G., Beauvillian, J.-C. and Masson, M., *Neurochem. Int.*, 1986, **8**, 327–338.
143. Porchet, M., Baert, J. L. and Dhainaut-Courtois, N., *Invert. Reprod. Develop.*, 1989, **16**, 53–62.
144. Porchet, M. and Dhainaut-Courtois, N., in *Neurohormones in Invertebrates* (eds Thorndyke, M. C. and Goldsworthy, G. J.), Cambridge University Press, Cambridge, 1988, pp. 219–234.
145. Kaiser, F., *Zool. Jahrb. Abt. Allg. Zool. Physiol.*, 1954, **65**, 59.
146. Gardner, C. R. and Walker, R. J., *Prog. Neurobiol.*, 1982, **18**, 81–95.
147. Al-Yousuf, S. A. in *International Review of Cytology* (eds Jeon, K. W. and Friedlander, M.), Academic Press, New York, 1992, pp. 231–308.
148. Prochet, M., Bulet, Ph., Curgy, J. J., Porchet-Hennere, E., Andries, J. C., Hesdin, A. D., Guissi-Kadri, S. and Smiri, Y., in *Neurobiology and Endocrinology of Selected Invertebrates* (eds Loughton, B. G. and Saleuddin, S. M.), Captus University Publications, Canada, 1990, pp. 27–46.
149. LeRoithe, D., Liotta, A. S., Roth, J., Schiloach, J., Lewis, M. E., Pert, C. B. and Krieger, D. T., *Proc. Natl. Acad. Sci. USA*, 1982, **79**, 2086–2090.
150. Roth, J., LeRoithe, D., Schiloach, J., Rosenzweig, J. L., Lesniak, M. A. and Havrankova, J., *New Engl. J. Med.*, 1982, **306**, 523–527.
151. Venturini, G., Carolei, A., Palladini, G. and Lauro, M. G., *Comp. Biochem. Physiol.*, 1983, **C74**, 23–25.
152. Venturini, G., Carolei, A., Palladini, G. and Margotta, V., in *CRC Handbook of Comparative Opioid and Related Neuropeptide Mechanisms* (ed. Stefano, G. B.), CRC Press, Boca Raton, Florida, 1986, vol. I, pp. 245–254.
153. Romero, R., Bueno, D., Sumoy, L., Burgaya, F., Soriano, M. A. and Baguna, J., *Gen. Comp. Endocrinol.*, 1991, **82**, 294.
154. Rémy, C. and Brossard, D., in *CRC Handbook of Comparative Opioid and Related Neuropeptide Mechanisms* (ed. Stefano, G. B.), CRC Press, Boca Raton, Florida, 1986, vol. I, pp. 155–164.
155. Colon-Urban, R., Rosales, N., Moradi, B. and Ocasio, V., in *CRC Handbook of Comparative Opioid and Related Neuropeptide Mechanisms* (ed. Stefano, G. B.), CRC Press, Boca Raton, Florida, 1986, vol. II, pp. 249–258.
156. Engbretson, G. A. and Chamberlain, S. C., *Invest. Opthalmol. Vis. Sci.*, 1981, **20** (Suppl.) 75.

Received 3 October 1994; revised accepted 21 July 1995

Carbon dioxide and nitrous oxide in the North Indian Ocean

M. Dileep Kumar, S. W. A. Naqvi, D. A. Jayakumar, M. D. George, P. V. Narvekar and S. N. de Sousa

National Institute of Oceanography, Dona Paula, Goa 403 004, India

The understanding of biogeochemical cycling of carbon dioxide and nitrous oxide in the oceans is essential for predicting the fate of anthropogenically emitted components. The North Indian Ocean, with its diverse regimes, provides us with a natural laboratory that can unravel the mechanisms controlling these gases with implications for the global aquatic bodies. In this review we discuss the anthropogenically impinged global budgets for these gases, summarize our results, largely collected under the Global Change Programme from the North Indian Ocean and evaluate the contributions from this region to the global sea-to-air fluxes.

OVER the geological times the earth system has been at its steady course towards attaining a steady state with respect to quantities and fluxes of materials in and among various reservoirs (i.e. land, ocean and atmosphere). Man's efforts for the betterment of his living standards have led to an unprecedented interference in the natural processes of this system in recent times. The localized pollution on land or in water bodies affects the biogeochemical processes largely locally, whereas the pollutants released to the atmosphere would be of global significance.

The rising levels of gases such as carbon dioxide (CO_2), nitrous oxide (N_2O), methane (CH_4) and freons (chlorofluorocarbons) in the atmosphere have been recognized not only to change the natural balances of carbon and nitrogen but also affect climate and human health[1]. The accumulation of these gases in the atmosphere enhances the greenhouse effect by absorbing the infrared radiation and leads to stratospheric ozone depletion. Nitrous oxide is 200 times more effective as a greenhouse gas, on per mole basis, compared to CO_2, but, due to a much lower concentration, it accounts for only about 5% of the total warming effect[2]. Under the International Geosphere–Biosphere Programme (IGBP) 'A study of global change', a lot of studies are being carried out globally on biogeochemical processes involving the greenhouse gases. We provide here a synthesis of the available data on CO_2 and N_2O cycling in the North Indian Ocean, largely generated as a result of our participation in the IGBP, in the context of global biogeochemical cycles.

Global balance of impinged CO_2 and N_2O

The atmospheric CO_2 and N_2O concentrations are increasing at annual rates[2] of 0.5% and 0.2%, respectively. The major anthropogenic activities contributing to the observed increases are fossil fuel combustion, use of fertilizers and deforestation. Of the total CO_2 released (7.3×10^3 Mt C yr^{-1}), 3.4×10^3 Mt C yr^{-1} is accumulating in the atmosphere whereas 2×10^3 Mt C yr^{-1} is being removed by the oceans[3] (Figure 1 *a*). However, there is considerable controversy regarding the amount associated with deforestation (1.9×10^3 Mt C yr^{-1}), which at present is attributed to an unknown sink or unaccounted recycled carbon accumulation in soils[3]. The increase in carbon fixation on land, due to the use of nitrogen fertilizers, has not so far been included in the carbon budget. It may account for the missing carbon to some extent. For instance, the agricultural nitrogen fixation in 1974 is estimated[4] to be about 39 Mt N yr^{-1}; this accounts for 0.65×10^3 Mt C yr^{-1}, with a C:N ratio of 19:1 for terrestrial plant material. Thus, this increased carbon fixation could roughly account for at least 34% of the presently missing carbon.

The land-to-air fluxes of N_2O (Figure 1 *b*) are essentially closer to the lower ranges given by Yoshinari[5] (cultivated soils, fertilized soils, unfertilized soils, tropical forests and temperate forests account for 2.0, 3.0, 4.5, 5.8 and 3.0 Mt N yr^{-1}, respectively), to which the flux (0.03 Mt N yr^{-1}) from watershed regions[6] was added in this attempt. The fossil fuel combustion and

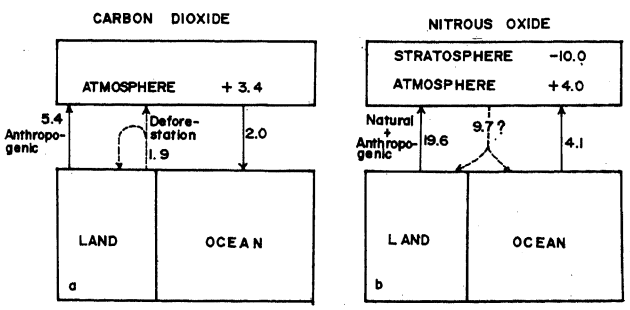


Figure 1. Global budgets for: *a*, anthropogenic component (only) of CO_2; and *b*, human-influenced total N_2O. CO_2 figures are in Gt C yr^{-1} (10^{15} g C yr^{-1}) and those of N_2O are in Mt N yr^{-1} (10^{12} g N yr^{-1}).

biomass burning contribute[7] 0.5 and 0.3 Mt N yr^{-1}, respectively, wherein nylon production results in an emission[8] of 0.42 Mt N yr^{-1}. We evaluated the fluxes from coastal wetlands and marshes to be 0.1 Mt N yr^{-1} based on the data of de Angelis and Gordon[9]. The upwelling zones gives a flux[10] of 1.0 Mt N yr^{-1} while the open ocean gives at least 3.0 Mt N yr^{-1}, which is closer to that given by Yoshinari[5], though the values in the literature vary over a wide range (1–44.5 Mt N yr^{-1}). The only known sink for N_2O is its photolytic decomposition in the stratosphere (Figure 1 *b*). After accounting for its stratospheric loss and atmospheric accumulation, we find that about 9.7 Mt N yr^{-1} is not balanced (Figure 1 *b*). This reveals that either the fluxes, especially from land, are overestimated or there could be an unknown sink, as in the case of carbon, but which is unlikely.

The unique North Indian Ocean

The presence of the Asian landmass, which bounds the Indian Ocean at relatively low latitudes, makes the North Indian Ocean geographically distinct. The Indian peninsula bifurcates the North Indian Ocean into two hydrographically and climatologically different water bodies – the Arabian Sea and the Bay of Bengal. The Arabia Sea occupies an area of 6.225×10^6 km^2 while the Bay of Bengal covers 4.087×10^6 km^2, excluding the Laccadive and Andaman seas. Seven major rivers flow into the Bay of Bengal (Brahmaputra, Ganges, Irrawady, Godavari, Mahanadi, Krishna and Kaveri) whereas three pour into the Arabian Sea (Indus, Narmada and Tapti). These rivers carry 1625×10^9 and 301×10^9 m^3 yr^{-1} of fresh water and 1387×10^6 and 195×10^6 tons yr^{-1} of suspended solids, respectively, into the Bay of Bengal and the Arabian Sea[11]. In addition to the smaller run-off, excess evaporation in the Arabian Sea leads to a negative water balance[12].

The North Indian Ocean is among the more productive parts of the oceans with a surface production (upper 1 m layer) amounting to 112 Mt C yr^{-1} (ref. 13). The productivity is particularly high in the Gulf of Oman and off the coasts of Arabia, Somalia and Southwest India in the Arabian Sea and off the Southeast Indian Coast, Visakhapatnam, Sri Lanka and Burmese coasts in the Bay of Bengal. An anomalous feature of the oceanography of the Indian Ocean is that the zones of intense upwelling occur not along the eastern boundary, as happens in the Pacific and the Atlantic Oceans, but in the northwestern parts off the Somali and the Arabia coasts. Consequently, the column productivity in the Arabian Sea is much higher compared to that in the Bay of Bengal. The high productivity causes intense oxygen demand in the subsurface layers. Due to this, the water column in the Arabian Sea experiences severe oxygen depletion, resulting in the onset of reducing conditions, thus making the North Indian Ocean one of the most important denitrification sites[14–18] in the world.

Carbon dioxide in the North Indian Ocean

A higher rate of biological production in the Arabian Sea compared to that in the Bay of Bengal coupled with the different hydrographies of the two regions causes vast differences in the biogeochemical cycling in subsurface waters. Although the entire North Indian Ocean experiences a severe oxygen depletion at mid-depths, the oxygen minimum layer is not only much more thicker in the Arabian Sea but is also far more intense. Consequently, while the Arabian Sea experiences denitrification within a large volume of intermediate waters, the Bay of Bengal is not an active water column denitrification site[18]. Denitrification is the most intense in the northeastern and central parts of the Arabian Sea. An intermediate particle maximum (as inferred from the vertical profiles of the beam attenuation coefficient,

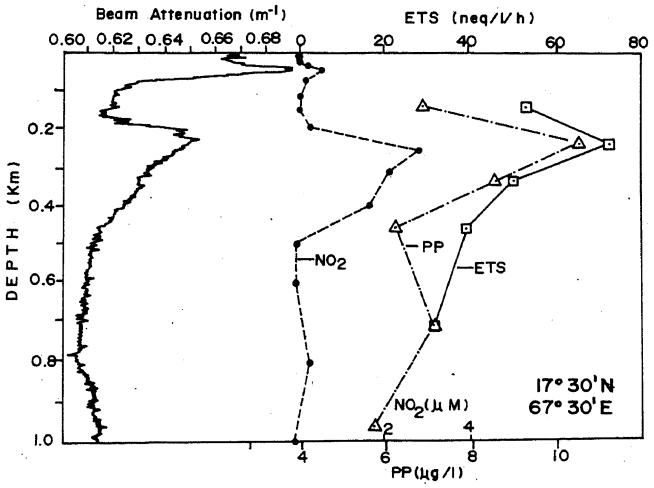


Figure 2. Profiles of beam attenuation coefficient (BAC), nitrite (NO_2^-), particulate protein (PP) and electron transport system (ETS) activities at a station in the central Arabian Sea.

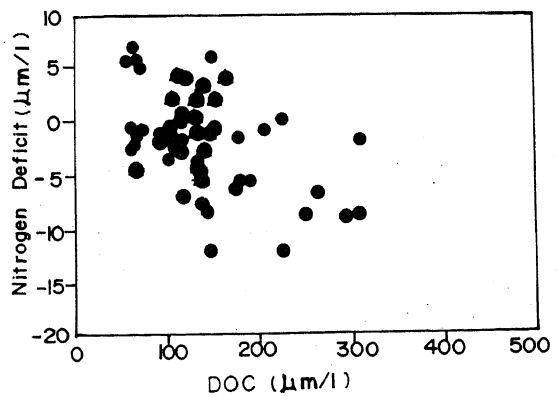


Figure 3. Relation between dissolved organic carbon (DOC) and nitrate deficit in the Arabian Sea.

BAC) has been found[19] to be associated with denitrification in the Arabian sea; particulate protein (PP) and activity of the respiratory electron transport system (ETS) also showed maxima (Figure 2; data collected during February 1992 on FORV Sagar Sampada Cruise 98) within the denitrifying layer. The raining particulate organic carbon from the overlying waters seems to be insufficient to support the measured metabolic rates, thus underscoring the role of organic carbon supplied through lateral processes in fuelling subsurface respiration within the oligotrophic central regions of the Arabian Sea[19]. This view is augmented by the observed negative relation[20] between DOC and inorganic nitrate deficits (Figure 3). The DOC levels are lower in the north than in southern Arabian Sea probably owing to the faster utilization and turnover of this form of carbon. Consequently, the total carbon dioxide (TCO_2) is also found to be inversely related to the DOC in the Arabian Sea[20].

In the subsurface layers the TCO_2 is higher in the North than in the South Arabian Sea because of the increased respiration rates in the northern region and also due to the aging of waters as they move northward. In the Bay of Bengal, on the other hand, there occur negligible north–south gradients in TCO_2 and O_2, presumably due to a much lower rate of organic decomposition within the water column from a combination of lower biological productivity and rapid settlement of organic matter through the water column, consequent to its incorporation into rapidly sinking terrigenous material. The increase in TCO_2 with depth is due to the decomposition of organic and inorganic materials. Surprisingly, in spite of the differences in productivity between the Arabian Sea and the Bay of Bengal and the rapid settlement of organic matter in the latter region, the levels of TCO_2 are almost the same (Figure 4 *a*,*b*). Sinking fluxes of organic matter ranged between 2.04 and 3.59 g C m^{-2} yr^{-1} in the Bay of Bengal[21], while they varied from 1.53 to 1.80 g C m^{-2} yr^{-1} in the Arabian Sea[22]. These differences, due to organic matter decomposition, could be balanced by increased contribution of skeletal dissolution to TCO_2 in the Bay of Bengal deep waters. Approximately 40–50% of the TCO_2 difference (at depths >1 km) between latitudes 10° and ~20°N in the Arabian Sea is accounted for by skeletal dissolution, but this mechanism is responsible for ~85% of TCO_2 increase between latitudes 11° and 18°N in the Bay of Bengal[23]. The prevailing reducing conditions lower the pH levels in the Arabian Sea as compared to those in the Bay of Bengal. As a result, the intermediate waters in the former region are more corrosive to skeletal carbonate materials than those in the latter[23]. Interestingly, skeletal dissolution is relatively more in the north and eastern parts than in the south and western regions of the Arabian Sea[24].

The pH within the oxygen minimum zone is generally lower than in other layers and hence the carbonate

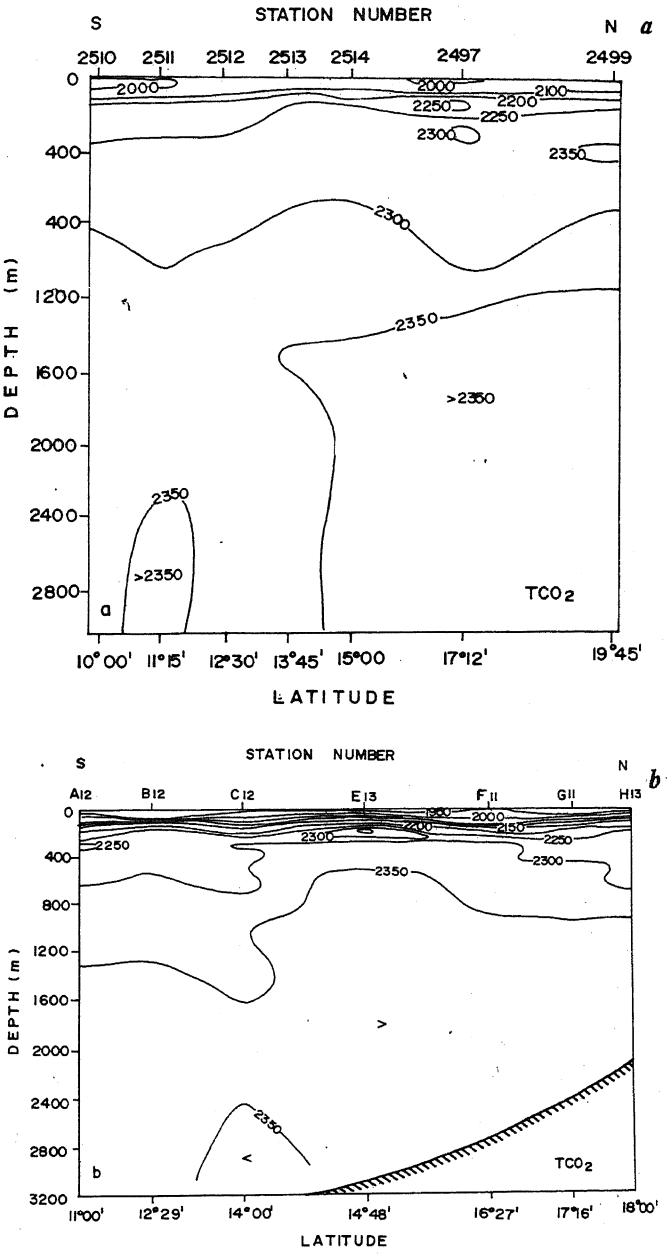


Figure 4. Distribution of total CO_2 (μM) in: *a*, Arabian Sea; *b*, Bay of Bengal.

equilibrium tends to shift relatively more towards the formation of carbonic acid. This increases the partial pressure of CO_2 (pCO_2) in this zone, although the TCO_2 concentrations are higher in the deep and bottom waters than in the oxygen-minimum zones (Figure 5 *a*,*b*). As the pH is considerably lower in the Arabian Sea than in the Bay of Bengal, the computed pCO_2, based on pH and TCO_2, is relatively higher in the former region[23]. A maximum pCO_2 of 1200 μatm occurs in the oxygen-minimum zone of the central Arabian Sea while it is 1100 μatm in the central Bay of Bengal. The pCO_2 is higher by about 100 μatm at any depth in the Arabian Sea than in the Bay of Bengal.

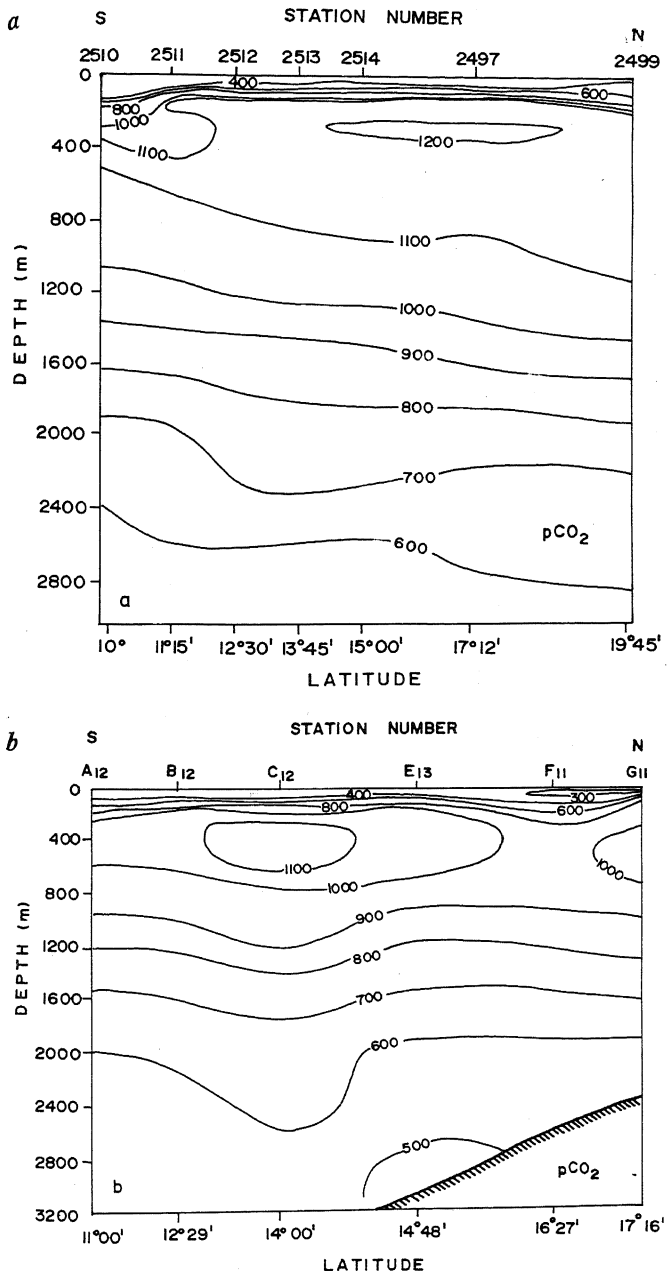


Figure 5. Distribution of pCO_2 (μatm) in: *a*, The Arabian Sea; *b*, Bay of Bengal.

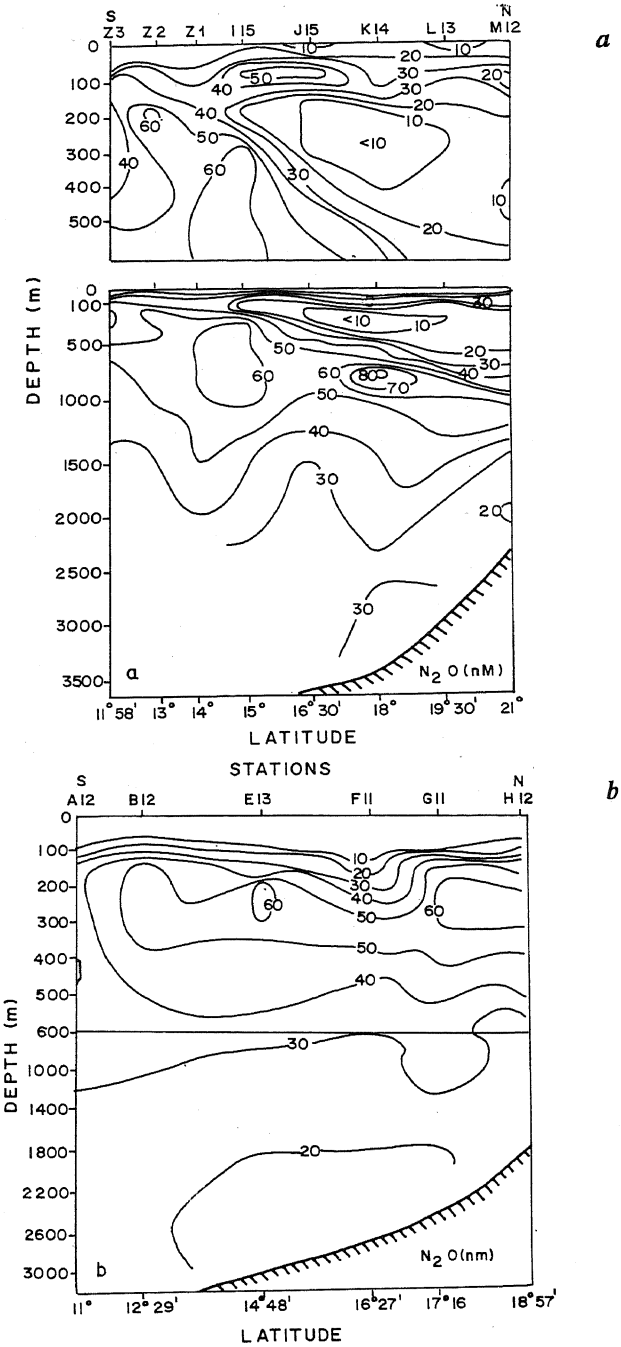


Figure 6. Nitrous oxide (nM) distribution in: *a*, The Arabian Sea; *b*, Bay of Bengal.

Nitrous oxide in the North Indian Ocean

Vertical profiles of N_2O generally exhibit maxima within the oxygen-minimum zones. The most significant point revealed by Figure 6 (data from ref. 25) is the occurrence of double maxima in N_2O with a minimum in the intensely reducing intermediate waters of the north-eastern Arabian Sea[25,26]. Outside the denitrification zone, N_2O exhibits a single peak coinciding with the oxygen minimum[25,27]. Consumption of N_2O within the strongly reducing zones appears to produce the minimum in N_2O within the denitrifying zone of the north-eastern Arabian Sea, while the maxima in N_2O occur at the boundaries of this zone, particularly at the lower boundary with concentrations as high as 80 nM (Figure 6 *a*). The N_2O concentrations increase in the northward direction (Figure 6 *b*) in the Bay of Bengal, with a maximum of 80 nM (occurring to the west of the transect shown in the figure). Maximum in the north corresponds to oxygen minimum. Although very low levels of oxygen are observed in subsurface layers of the northernmost Bay of Bengal, the absence of a secondary nitrite shows that denitrification may not occur within the water column in this region. The occurrence of strong

sources and sinks of N_2O in close proximity to each other is expected to lead to a very rapid turnover of N_2O in this region, which may have important implications for the residence time and isotopic composition of oceanic N_2O. The relations between apparent N_2O production and the apparent oxygen utilization are positively linear in different depth ranges in the Arabian Sea[25] and the Bay of Bengal[27]. The slopes estimated for the Arabian Sea are higher than those for the other oceanic regions, implying that the rate of N_2O production in this region is much higher[25]. Based on the horizontal gradients and that expected from oxygen utilized, significant quantities of N_2O are expected to be consumed within the sediments[25,27].

The enhanced production of N_2O in the subsurface layers leads to large vertical gradients within the upper thermocline and the consequent large fluxes into the surface layers through eddy diffusion. The Arabian Sea surface waters are relatively more supersaturated (average of 186%; refs 25, 26) as compared to those in the Bay of Bengal (125%; ref. 27). This higher saturation is related to higher vertical diffusion coefficients in the Arabian Sea (0.55 $cm^2 s^{-1}$; ref. 25) than in the Bay of Bengal (0.16 $cm^2 s^{-1}$; ref. 27), in the thermocline region. The suppressed vertical mixing in the Bay of Bengal is caused by strong thermohaline stratification caused by river run-off.

Sea-to-air fluxes in the North Indian Ocean

The differences between the concentrations observed at sea surface and those expected from the solubility calculations, based on the abundance of these gases in the atmosphere, determine the magnitude and direction of gas fluxes between the atmosphere and the sea. The abundance of N_2O and pCO_2 in surface layers will be governed by the gradients in these properties below the surface mixed layer and the rate of vertical mixing. The levels of pCO_2 (Figure 5) and N_2O (Figure 6) are higher within the oxygen-minimum zones in the North Indian Ocean sustaining large upward fluxes of these gases across the thermocline.

The fraction of CO_2 in air, in equilibrium with surface waters, over the Arabian Sea is calculated to be >0.6, which increases northward[24]. This also indicates that the Arabian Sea acts as a source of CO_2 to the atmosphere, in accordance with the earlier predictions[28–30]. The higher pCO_2 gradient between the deep and surface layers in the Arabian Sea than that in the Bay of Bengal (Figure 5) results in higher average exchange rates in the former region (2.0 mmol $m^{-2} d^{-1}$) than in the latter (0.17 mmol $m^{-2} d^{-1}$); here an atmospheric CO_2 of 350 µatm was used. Within the Arabian Sea the fluxes are higher in the north than in the south. The overall atmospheric fluxes of CO_2 are much larger from the Arabian Sea (55 Mt C yr^{-1}) compared to those from the Bay of Bengal (3 Mt C yr^{-1}). This difference arises from a combination of lower pCO_2 gradient, lesser vertical diffusion coefficients and smaller area of the Bay of Bengal.

A high degree of supersaturation of N_2O in the surface waters of the North Indian Ocean results in significant evasion of this gas to the atmosphere. As in the case of CO_2, the N_2O in surface layers is also controlled by upward diffusion from the subsurface layers. The average emission rates for this gas are 4.46 µmol $m^{-2} d^{-1}$ in the Arabian Sea and 0.65 µmol $m^{-2} d^{-1}$ in the Bay of Bengal. These rates amount to N_2O fluxes of 0.28 and 0.03 Mt N yr^{-1} from the Arabian Sea and the Bay of Bengal, respectively[25,27]. Hence, relatively large amounts of these greenhouse gases are emitted from the former region than from the latter.

Budgets for CO_2 and N_2O in the North Indian Ocean

In order to understand the biogeochemical cycling of an element it is essential to quantify its fluxes between various major reservoirs as well as between various compartments within a reservoir. The situation in the

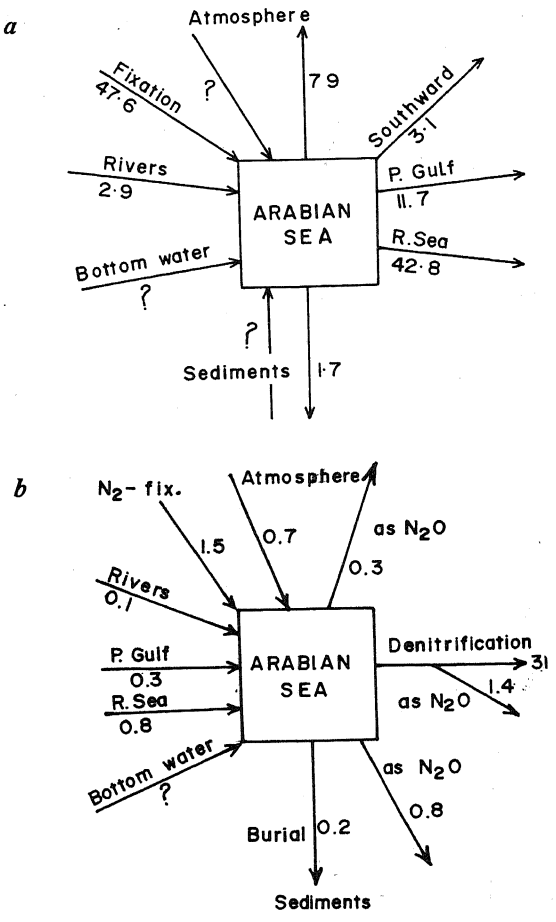


Figure 7. Budgets for: *a*, carbon; *b*, nitrogen (including N_2O) in the Arabian Sea.

Indian Ocean is complicated by the presence of marginal water bodies that continuously exchange water and solutes with the Arabian Sea and with river run-off that brings in enormous amounts of inorganic and organic substances into the Bay of Bengal. Recently, attempts have been made to quantify the fluxes between various reservoirs in the Arabian Sea[30,31]. Figure 7 depicts the estimated net carbon and nitrogen fluxes into and out of the Arabian Sea. The flow of carbon into the Arabian Sea from the rivers is the only estimated external input (using an annual discharge rate of 238 km^3 and an average HCO_3^- concentration[32] of 12 mg C l^{-1}), which is close to that (2.3 Mt C yr^{-1}) evaluated earlier[30]. The unknowns are its influx from the atmosphere, inflow from the south of equator and upward diffusion of regenerated CO_2 from the sediments. The biologically fixed carbon transported to subsurface Arabian Sea layers, through sinking of particulate organic material, is estimated[31,33] to be about 47.6 Mt C yr^{-1}. The estimated fluxes (Figure 7) reveal that the Arabian Sea serves as a source of CO_2 to the atmosphere and the surrounding bodies, the Persian Gulf and the Red Sea. Carbon is also supplied to the underlying sediments and to some extent to the South Indian Ocean. These estimates indicate that the Arabian Sea is constrained by a CO_2 deficit of ~88 Mt C yr^{-1}. In order to maintain the CO_2 in the Arabian Sea at steady state, this amount must be supplied from the south, through the inflowing bottom water, and from the atmosphere.

As our state of understanding of the processes regulating N_2O in the Arabian Sea is still rudimentary, it is difficult to work out the budget for this gas. Figure 7 shows the net fluxes of N_2O together with those of nitrogen[25,30,31,33]. In contrast to their negative contribution to CO_2, the Persian Gulf and the Red Sea act as net sources of nitrogen to the Arabian Sea. The losses of nitrous oxide from the Arabian Sea amount to 2.5 Mt N yr^{-1}. Hence, the Arabian Sea should produce an equal amount annually.

The estimated carbon fluxes into the confluence zones of the rivers Ganges and Brahmaputra are 14.6 Mt C yr^{-1} and 17.1 Mt C yr^{-1}, respectively[34]. The carbon input through dissolved and particulate organic forms from the Ganges amounts[34,35] to 1.5–2.0 Mt C yr^{-1}. The Bay of Bengal serves as a mild source of atmospheric N_2O and CO_2 during the southwest premonsoon[23,27] but as a sink for CO_2 during the northeast monsoon[36].

Contribution of the North Indian Ocean to global air–sea fluxes

The contribution of the North Indian Ocean to the global sea-to-air flux[2] of CO_2 of 90 Gt C yr^{-1} appears to be small (0.06 Gt C yr^{-1}). However, it amounts to about 5% of the efflux from global equatorial oceans. On the contrary, the N_2O emission from this region is globally significant as it is about 8% of the estimated atmospheric input from the oceans. However, the region under study is significantly influenced by the highly variable monsoonal winds and intermittent cyclones. The transfer velocities of gases across the air–sea interface increases with increase in wind velocity. These transfer velocities have been computed to be enhanced seven times during the monsoon season compared to those during the non-monsoon seasons[37]. However, all the estimates made so far, including those for global fluxes, are derived from observations made during the calm seasons and do not consider the stability of sea surface as well as contributions from bubbles, and, hence, should be considered as underestimates. Further, the Arabian Sea experiences strong upwelling seasonally associated with the monsoons. Hence, the dynamics of these gases need to be studied in detail in order to make more accurate estimates of sea-to-air fluxes from this most dynamic region of the world oceans.

1. IGBP, Report No. 1, ICSU, Stockholm, 1986.
2. Rodhe, H., *Science*, 1990, **248**, 1217–1219.
3. Siegenthaler, U. and Sarmiento, J. L., *Nature*, 1993, **365**, 119–125.
4. Pratt, P. F., Barber, J. C., Corrin, M. L., Goering, J., Hauck, R. D., Johnston, H. S., Klute, A., Knowless, R., Nelson, D. W., Pickett, R. C. and Stephens, E. R., *Clim. Change*, 1977, **1**, 109–135.
5. Yoshinari, T., in *Denitrification in Soil and Sediment* (eds Revsbech, N. P. and Sorensen, J.), Plenum Press, New York, 1990, pp. 129–150.
6. Bowden, W. B. and Bermann, F. H., *Science*, 1986, **233**, 867–869.
7. Codispoti, L. A., Elkins, J. W., Yoshinari, T., Friederich, G. E., Sakamoto, C. M. and Packard, T. T., in *Oceanography of the Indian Ocean* (ed. Desai, B. N.), Oxford & IBH, New Delhi, 1992, pp. 271–284.
8. Thiemens, M. H. and Trogler, W. C., *Science*, 1991, **251**, 932–934.
9. de Angelis, M. A. and Gordon, L. I., *Estuarine Coastal Shelf. Sci.*, 1985, **20**, 375–386.
10. Codispoti, L. A. and Christensen, J. P., *Mar. Chem.*, 1985, **16**, 277–300.
11. Subramanian, V., *Curr. Sci.*, 1993, **64**, 928–930.
12. Ramanathan, K. R. and Pisharoty, P. R., in *World Water Balance*, IASH–UNESCO–WMO, Gentbrugge, 1972, vol. I, pp. 39–41.
13. Qasim, S. Z., *Indian J. Mar. Sci.*, 1977, **6**, 122–137.
14. Sen Gupta, R., Rajagopal, M. D. and Qasim, S. Z., *Indian J. Mar. Sci.*, 1976, **5**, 201–211.
15. Sen Gupta, R. and Naqvi, S. W. A., *Deep-Sea Res.*, 1984, **31**, 671–706.
16. Naqvi, S. W. A., *J. Mar. Res.*, 1987, **45**, 1049–1072.
17. Naqvi, S. W. A., *Oceanol. Acta*, 1991, **14**, 281–290.
18. Rao, C. K., Naqvi, S. W. A., Kumar, M. D., Varaprasad, S. J. D., Jayakumar, D. A., George, M. D. and Singbal, S. Y. S., *Mar. Chem.*, 1994, **47**, 279–290.
19. Naqvi, S. W. A., Kumar, M. D., Narvekar, P. V., de Souza, S. N., George, M. D. and D'Silva, C., *J. Geophys. Res.*, 1993, **98**, 16469–16479.
20. Kumar, M. D., Rajendran, A., Somasundar, K., Haake, B., Jenisch, A., Shuo, Z., Ittekkot, V. and Desai, B. N., *Mar. Chem.*, 1990, **31**, 299–316.
21. Ittekkot, V., Nair, R. R., Honjo, S., Ramaswamy, V., Bartsch, M., Manganini, S. and Desai, B. N., *Nature*, **351**, 385–387.

22. Nair, R. R., Ittekkot, V., Manganini, S., Ramaswamy, V., Haake, B., Degens, E. T., Desai, B. N. and Honjo, S., *Nature*, 1989, **338**, 749–751.
23. George, M. D., Kumar, M. D., Naqvi, S. W. A., Banerjee, S., Narvekar, P. V., de Souza, S. N. and Jayakumar, D. A., *Mar. Chem.*, 1994, **47**, 243–254.
24. Kumar, M. D., Rajendran, A., Somasundar, K., Ittekkot, V. and Desai, B. N., in *Oceanography of the Indian Ocean* (ed. Desai, B. N.), Oxford & IBH, New Delhi, 1992, pp. 313–325.
25. Naqvi, S. W. A. and Noronha, R. J., *Deep-Sea Res.*, 1991, **38**, 871–890.
26. Law, C. S. and Owens, N. J. P., *Nature*, 1990, **346**, 826–828.
27. Naqvi, S. W. A., Jayakumar, D. A., Nair, M., Kumar, M. D. and George, M. D., *Mar. Chem.*, 1994, **47**, 279–290.
28. Takahashi, T., *Oceanus*, 1989, **32**, 22–29.
29. Tans, P. P., Fung, I. Y. and Takahashi, T., *Science*, 1990, **247**, 1431–1438.
30. Somasundar, K. Rajendran, A., Kumar, M. D. and Sen Gupta, R., *Mar. Chem.*, 1990, **30**, 363–377.
31. Naqvi, S. W. A., Sen Gupta, R. and Kumar, M. D., in *Interactions between Global Climate Subsystems, The Legacy of Hann*, Geophysical Monograph No. 75, IUGG Volume 15, American Geophysical Union, 1992, pp. 85–92.
32. Pande, K., Sarin, M. M., Trivedi, J. R., Krishnaswami, S. and Sharma K. K., *Chem. Geol.*, **116**, 245–259.
33. Naqvi, S. W. A., Noronha, R. J., Shailaja, M. S., Somasundar, K. and Sen Gupta, R., in *Oceanography of the Indian Ocean* (ed. Desai, B. N.), Oxford & IBH, New Delhi, 1992, pp. 285–311.
34. Safiullah, S., Mofizuddin, M., Ali, S. M. I. and Kabir, S. E., in *Transport of Carbon and Minerals in Major World Rivers*, SCOPE/UNEP Sonderband Heft 64, 1987, Part 4, pp. 435–442.
35. Kumar, M. D., George, M. D. and Sen Gupta, R., in *Oceanography of the Indian Ocean* (ed. Desai, B. N.), Oxford & IBH, New Delhi, 1992, pp. 347–358.
36. Kumar, M. D., Naqvi, S. W. A., George, M. D. and Jayakumar, D. A., submitted to *J. Geophys. Res.*
37. Etcheto, J., Boutin, J. and Merlivat, M., *Tellus*, 1991, **43**, 247–255.

ACKNOWLEDGEMENTS. We thank the Director, NIO, Dr A. P. Mitra and Dr R. Sen Gupta for their encouragement and support.

Received 2 September 1994; revised accepted 8 September 1995

RESEARCH ARTICLES

Internal waves – A novel measurement technique

Rao Tatavarti, P. N. Ananth, K. Rajasree*, V. Vidyalal*, P. Radhakrishnan*, V. P. N. Nampoori* and C. P. G. Vallabhan*

N. P. O. L., Defence Research and Development Organization, Thrikkakara, Cochin 682 021, India
*Laser Division, International School of Photonics, Cochin University of Science and Technology, Cochin 682 022, India

We present here a novel opto-electronic technique to directly measure the internal wave oscillations, created by systematically disturbing a density stratified water column with a rotating propeller in a laboratory wave tank. The technique, based on the principle of laser beam deflection (LBD), apart from being noninvasive and highly sensitive, does not have the customary constraints on the sampling frequency and the duration of data sampling, as is the case with conventional experimental procedures. The noninvasive nature of this technique has good potential in future studies of diffusion and mixing processes in different disciplines.

INTERNAL waves in the ocean are ubiquitous, being associated with wind and air pressure fluctuations, changes in ocean floor topography, wakes of moving bodies and instabilities in the water body. Studies on internal waves enable a better understanding of the various mixing processes in the ocean, which result in the redistribution of momentum, temperature (heat) and salinity, responsible for climatic changes and, nutrients responsible for biological productivity. The oceanic internal waves occupy a vast continuum of spatial and temporal scales[1]. Available data on oceanic motions suggest horizontal scales ranging from a few tens of meters to a few kilometers and temporal scales ranging from the inertial period f^{-1} (order of magnitude in days), to the Brunt Vaisala period N^{-1} (order of magnitude in minutes).

Internal wave measurements

To the best of our knowledge, only indirect eulerian field of internal waves were measured by monitoring a closely related physical parameter like temperature or current velocity. Lagrangian field measurements of internal waves involve neutrally buoyant floats that are tracked, or sensor arrays towed horizontally through water. The literature[2–6] points out the limitations of the indirect eulerian measurements, whereas the established Lagrangian techniques are constrained by the uncertainties in understanding the underlying physics[7].

The electrical conductivity of salt water in the ocean was earlier considered to provide a convenient way of

measuring density in a stratified fluid. In this technique, the total impedance between a relatively large fixed electrode and a much smaller, movable platinum electrode is measured in an A.C. bridge network. However, two difficulties are encountered in practice. First, the probe impedance tends to drift with time. Second, time response of the probe is limited, roughly by the time required to convect material across the probe volume[8].

Remote sensing of oceanic internal waves from airborne synthetic aperture radar (SAR) imagery, a state-of-art measurement technique, involves intricate calibrations and signature extraction techniques. As the patterns on SAR imagery caused by mesoscale oceanic features are only a surface manifestation of subsurface physical processes, it is imperative to develop a robust theoretical model which links the subsurface processes to SAR-observed surface patterns. Both the manner in which SAR data are calibrated and the techniques used to extract the signatures associated with the surface patterns of interest can significantly influence the magnitude and structure of the SAR signature[9].

Laboratory investigations of internal waves primarily utilize noninvasive flow visualization techniques, as *in situ* monitoring of the complex internal wave dynamics is difficult[8]. Though flow visualization techniques give a comprehensive picture of the processes involved, they are primarily qualitative in nature. Flow visualization techniques are broadly subdivided into those that make use of light scattered by tiny particles in the fluid and those that make use of variations in the refractive index. Among the former, laser Doppler anemometry is now a standard means of obtaining fluid velocities. The fluid velocity can be measured with a high degree of accuracy as a function of time but only at a single point in the fluid at any given time[10]. Although, recent studies have shown that the instantaneous velocity field over a complete plane of interest in the fluid is now possible with speckle photography, the potential applications to fluid mechanics have not yet been fully explored.

Of the methods making use of refractive index variations, the laser beam deflection (LBD) or the mirage effect has been increasingly used in recent years for various applications[11,12], in which, the optical probe beam deflection due to an *artificially* generated refractive index gradient (RIG) is generally monitored.

Novel opto-electronic technique to monitor internal waves

Internal waves in a stratified fluid involve changes of refractive index across the field to be investigated, which may be visualized or photographed using optical methods that depend on the effects of the refractive index changes on the transmission of light. Several techniques are available of which the Schlieren techniques have been extensively used. The main use of Schlieren and direct shadow methods is to show the position and shapes of regions of density changes, such as those occurring in high speed flows, mixing problems and problems of free and forced convection.

Concept

In this section we present a novel optical technique, based on the principle of LBD, to directly quantify internal waves in the laboratory, generated by the rotating propeller of a self-propelled body submerged in a stratified water column.

In an inhomogeneous medium as in the presence of internal waves in a stratified water column, a laser beam gets deflected because of the *naturally* occurring, time-varying changes in the density of the medium (and therefore, the RIG) at a particular location. For laser beams with a gaussian intensity profile, the output voltage change monitored by a small area photo detector system is closely related to the changes in beam deflection[13]. Using this principle we were able to directly monitor internal waves generated by the rotating propeller of a stationary body placed inside a 6 mm thick perspex tank (1.8 m × 0.3 m × 0.5 m) containing an exponentially stratified salt water column.

The technique depends on the deflection of a ray of light from its undisturbed path when it passes through a medium in which there is a component of the gradient of the refractive index normal to the ray. The deflection of light ray is a measure of the first derivative of density with respect to the distance, i.e. the density gradient and therefore, it can be shown that the curvature of the ray is proportional to the refractive index gradient normal to the ray. If the Y axis is taken in the direction of the undisturbed ray, the curvatures in the XY and ZY planes respectively are given by

$$\frac{\partial^2 x}{\partial y^2} = \frac{1}{\mu}\frac{\partial \mu}{\partial x}. \qquad (1)$$

$$\frac{\partial^2 z}{\partial y^2} = \frac{1}{\mu}\frac{\partial \mu}{\partial z}.$$

If the total angular deflections in the XY and ZY planes are denoted by ε'_x and ε'_z respectively, then

$$\varepsilon'_x = \int \frac{1}{\mu}\frac{\partial \mu}{\partial x}\mathrm{d}y, \qquad (2)$$

$$\varepsilon'_z = \int \frac{1}{\mu}\frac{\partial \mu}{\partial z}\mathrm{d}y.$$

If the optical disturbance is in the working section of the wave tank, the ray of light will be deflected on leaving the tank so that

$$\mu \sin \varepsilon' = \mu_0 \sin \varepsilon, \qquad (3)$$

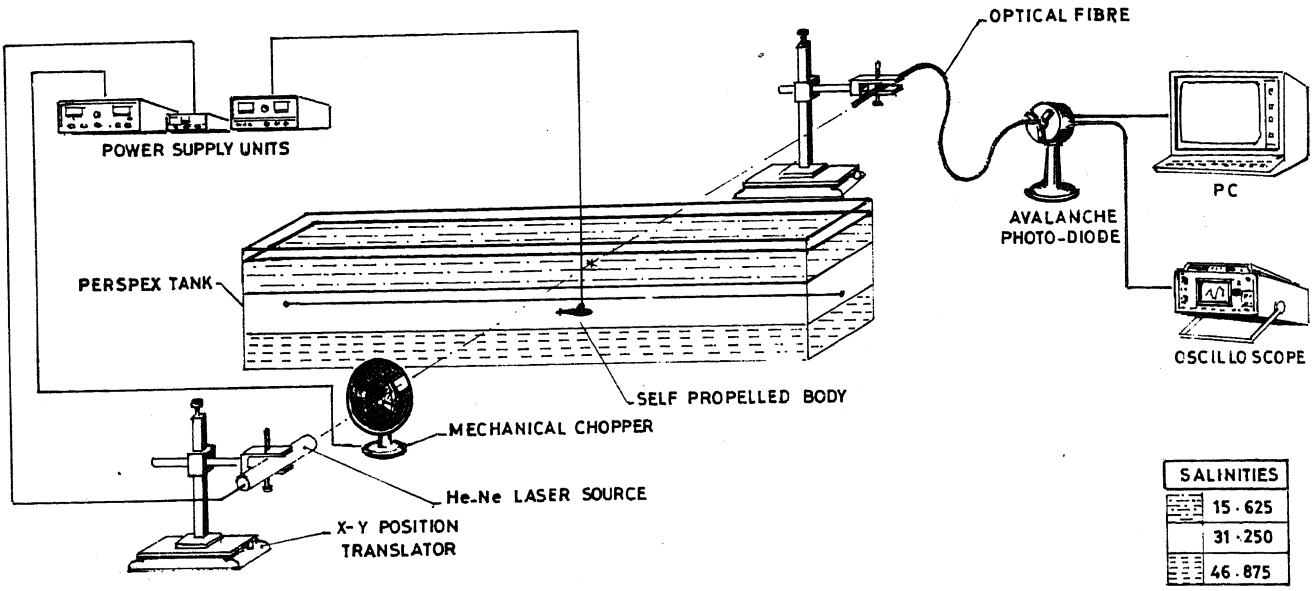


Figure 1. Schematic diagram of the experimental set-up to monitor internal waves in a stratified fluid employing the laser beam deflection technique.

where μ_0 is the refractive index of the air surrounding the tank, and μ is the refractive index in the working section. Thus, the final angular deflection ε measured beyond the tank is given by

$$\varepsilon_x = \frac{1}{\mu_0}\int \frac{\partial \mu}{\partial x}\,dy, \qquad \varepsilon_z = \frac{1}{\mu_0}\int \frac{\partial \mu}{\partial z}\,dy, \tag{4}$$

where the integrals are taken over the width of the working section. In the case of two-dimensional flow in a wave tank of width L these expressions simply become

$$\varepsilon_x = \frac{L}{\mu_0}\frac{\partial \mu}{\partial x}, \qquad \varepsilon_z = \frac{L}{\mu_0}\frac{\partial \mu}{\partial z}, \tag{5}$$

the deflections being in the direction of the RIG, i.e. towards the region of higher density.

The noninvasive nature of this technique has potential application in the study of diffusion and mixing processes in different disciplines, especially biofluid mechanics. The technique can be adapted for oceanographic field applications.

Laboratory studies

Experimental set-up. An intensity stabilized 5 mw, He–Ne continuous-wave laser beam (half power beam width of 0.84 mm) was used as the probe. The deflections of the probe beam caused by the induced RIG in the medium were detected by a position sensitive detector (PSD). The polished tip of a multimode optical fibre (80/120 μm) served as the sensor head of the PSD. The other end of the optical fibre was coupled to an avalanche photodiode (APD, *Thorn EMI* make), which was incorporated into an appropriate detector circuit of very low noise level. The output voltage from the APD circuit is piped into a personal computer using a high speed analogue to digital conversion data acquisition card. The resolution of the analogue to digital voltage conversion was 0.0024414 V. As the probe beam is continuous, the sampling frequency and duration of the data sample can be chosen without any constraints. The experimental set-up is shown in Figure 1.

Prior to each experiment, a mechanical chopper placed between the laser source and the tank was used to maximize the output signal by properly aligning the probe beam and the PSD. During the actual experiments, however, the chopper was removed. Initially, experiments were carried out to ascertain that temporally and spatially varying beam deflections occur only due to changes in the RIG of the medium in the wave tank, but not due to the RIG at the air–perspex–water interfaces. Hence, the time-varying beam deflections or changes in the output voltages can be attributed to the internal wave (density) fluctuations, generated by the propeller motions of the stationary body inside the stratified water column.

Calibrations. Utilizing known density samples in the perspex tank, the output voltages of the laser beam propagating through the samples were monitored. An output voltage of 3.62 V was observed for a sample density of 1.0043 g/cc and an output voltage of 1.08 V

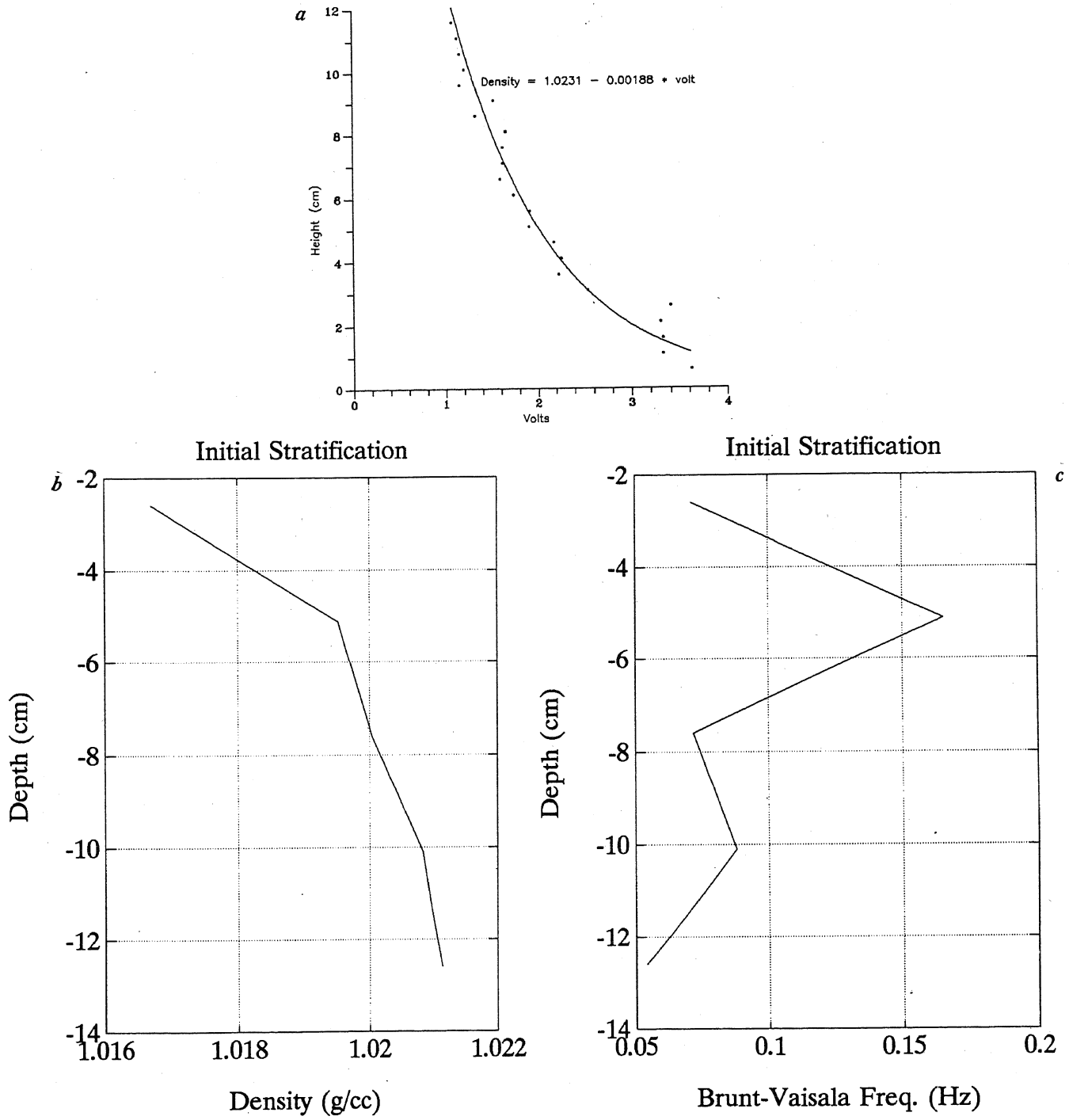


Figure 2. Initial output voltage distribution (*a*), calibrated density stratification (*b*), and the associated Brunt-Vaisala frequency [$N(z)$] profile (*c*), in the laboratory tank before the start of the experiment, where $N^2(z) = g/\rho_o\{\partial\rho/\partial z\}$.

was observed for a sample density of 1.0231 g/cc. Initially, a step-distribution was obtained by sequentially introducing the bottom layer (1.0231 g/cc), the middle layer (1.0113 g/cc) and the top layer (1.0043 g/cc) into the tank. Before introducing the middle and top layers, a wax paper was stretched on top of the water in the tank and the next layer added, after which the wax paper was rolled up from one end of the tank ensuring minimal disturbance. Molecular diffusion was allowed to take place for about a day, which resulted in the exponential distribution.

The output voltages of the laser beam propagating at different levels of the tank, in the vertical direction were monitored. Figure 2 *a* shows the vertical profile of output

voltages approximately a day after introducing the three layers in the tank. The exponential distribution is an artifact of the molecular diffusion. The calibration constants were obtained after solving the simultaneous (straight line) equations relating density and voltage. The gain and offset calibration constants were found to be 0.00188 and 1.0231 respectively. Figures 2 *b* and 2 *c* show the corresponding density and Brunt-Vaisala distributions in the tank. The output voltage from the APD circuitry which is piped into the PC as digital data is finally calibrated as density fluctuations.

Experiments. At the start of each experimental run the propeller was switched on. The sampling frequency was chosen as 100 Hz. In order to avoid reflections from the end walls of the tank, the duration of sampling for each run was limited to 2 min, the monitored travel times of hydrodynamical signatures from source to the end wall being greater than 2 min. Four major experiments covering 80 data runs in total were conducted with the laser source-detector system and the stationary body with its rotating propeller placed at different positions, for the same stratification. Detailed studies on different aspects of internal wave dynamics and mixing are being reported elsewhere. In this article, however, we present data from one of the typical experiments. During this experiment, time series observations of density fluctuations were monitored at 24 different locations along the wall of the tank forming a vertical array of time series observations as shown in Figure 3. In this article, however, data from the first vertical array of observations are only being reported. For all the observations, the position of the stationary body with its rotating propeller was fixed and only the position of the laser source-detector system was varied. Before each observational run, the propeller was switched on and after each observational run of 2 min duration, the propeller was switched off, allowing the water to settle down to its equilibrium position. As the forcing imparted to the water inside the tank was not strong (just sufficient to create internal wave fluctuations only) it is assumed that the mixing created by propulsion of the stationary body does not alter the characteristics of the stratified water column in the tank in a significant manner. The wake generated by a submerged body moving through a density-stratified medium is a turbulent region of mixed fluid of homogeneous density. The observations of earlier researchers that due to the stabilizing effect of the stratification, turbulence near the wake boundary is rapidly damped and does not cause appreciable mixing of the fluid with its surroundings, give ample support to our assumption. Therefore, as the mixed fluid seeks to return to its equivalent density level, the region collapses vertically and internal waves are consequently generated. At each location, the mean density value obtained during calibration experiment was deducted from the time series observations at that location to ensure that the observations reflect the actual density fluctuations caused by internal wave dynamics.

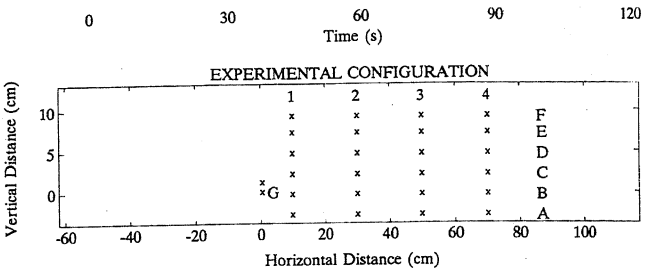


Figure 3. Experimental configuration. The crosses denote the exact locations where measurements were made. The *X*-direction is denoted by 1, 2, 3 and 4. The *Z*-direction is denoted by A, B, C, D, E and F. The propeller (origin) was lying 0.5 cm to the left of G.

Conventional Schlieren techniques – Limitations

Our technique is similar in principle to the Schlieren system[14,15]. However, the primary difference is that a conventional Schlieren system enables only qualitative flow visualization based on frame-to-frame observations, whereas our technique provides quantitative point-probe diagnostics continuously.

Conventional Schlieren techniques tend to become ineffective in monitoring smaller density changes of a medium. Indeed, when the density changes are very low (as in our study), conventional Schlieren methods can no longer be used as there is a progressive loss of contrast in the records that cannot be overcome by any normal modification to the apparatus or in the photographic processing.

In addition to limitations to the dimensions of mirrors and lenses used in conventional Schlieren techniques, the effects of diffraction may result in a change of illumination on the viewing screen, which differs appreciably in both extent and magnitude from that expected on the basis of geometrical optics. These effects may considerably affect the accuracy of the Schlieren technique when used to quantify density variations.

Even when high-speed cine cameras are used to quantify complicated dynamical processes (like internal waves in stratified fluid), the picture sharpness is usually less satisfactory in a direction along the film than across it[16].

Toepler–Schlieren photography can only offer the phase configurations, but our technique can determine the amplitude and phase of the signature using standard signal processing techniques. Moreover, the fact that none of the limitations associated with conventional Schlieren techniques arise when internal waves are monitored using the present method justifies that the technique reported here is more effective and sensitive.

Results and discussion

Typical time series plots of the density fluctuations, i.e. the raw data, along with a plot depicting the noise in

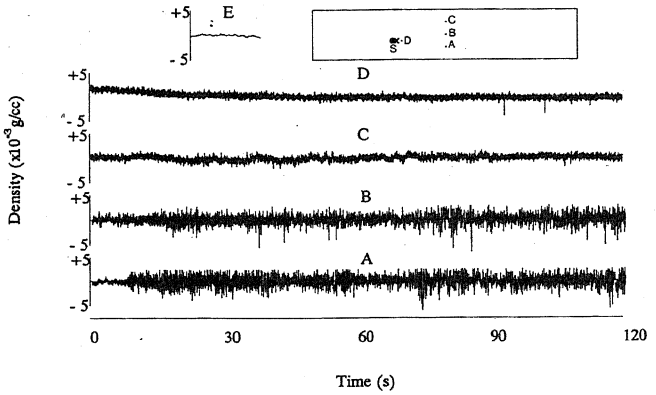


Figure 4. Time series of the observed density fluctuations caused by the turbulent wake and internal waves at different vertical locations A, B, C, D. Data run E depicts the ambient noise level at the location of A when the propeller was inoperative. Inset is a schematic diagram showing locations of runs A, B, C, D and the body S.

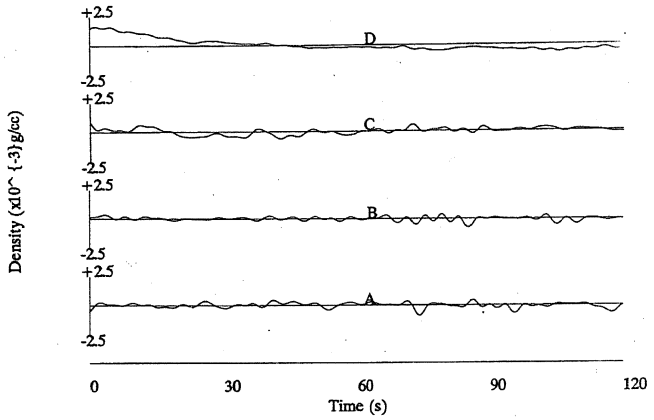


Figure 5. Time series of the low passed density fluctuations at different vertical locations. The data are the same as those shown in Figure 4.

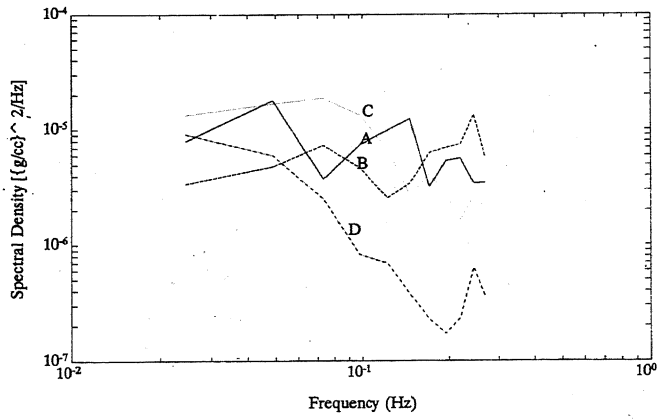


Figure 6. Spectral densities of the low passed density fluctuations recorded at different vertical locations A, B, C, D. The units on Y-axis are $[(g/cc)^2/Hz]$. Frequency resolution is 0.0244 Hz. Degrees of freedom ≅ 10.

measurements (i.e. when the signatures were observed while the water in the tank was undisturbed, the propeller being stationary) are shown in Figure 4. Two inferences become evident from Figure 4. One, the amplitudes of the fluctuations are small, indicating the minimal forcing applied on the propeller of the stationary body to avoid reflections and significant mixing; and two, the technique is very sensitive having minimal noise as can be seen in the uppermost plot of Figure 4. The maximum Brunt-Väisala frequency for our stratification was of the order of 0.2 Hz as reported earlier. Hence, the raw time series data were low passed using a Chebyshev filter with a cut-off frequency at 0.2 Hz to remove extraneous signatures, presumably associated with nonlinearities and turbulence, from analysis of internal wave dynamics. However, it may be noted that the detector does not yield any output corresponding to that part of the turbulence which does not induce any change in RIG in the medium. Figure 5 shows some typical time series plots of the low passed density fluctuations considered in the present study. Spectral density plots of the low passed raw data are shown in Figure 6. Spectral computations were performed using a Hanning window with 50% overlap. The spectral estimates have approximately 10 degrees of freedom with a frequency resolution of 0.0244 Hz. The spectral observations have been resolved into eight frequency bands. Significant spectral energy levels at different frequency bands can be observed in Figure 6.

Table 1. Wavenumber k_x (values in rad/cm)

	Horizontal distance (cm)					
	10 –30	30 –50	50 –70	10 –30	30 – 50	50 – 70
	Frequency = 0.0244 Hz			Frequency = 0.1465 Hz		
F	–0.030	0.147	–0.138	0.102	0.143	–0.044
E	–0.206	0.199	–0.232	0.211	–0.256	0.142
D	0.208	0.031	–0.267	0.239	–0.022	–0.095
C	–0.212	0.007	0.030	0.110	–0.098	0.109
B	–0.049	0.069	–0.018	0.008	–0.470	–0.045
A	–0.259	0.190	0.074	0.158	–0.071	0.067

The origin of the horizontal distances is the source of forcing. A – F denote vertical locations separated by 2.5 cm.

Tables 1 and 2 show the scalar wavenumbers, corresponding to two frequencies in the internal wave band, computed as $k_x = \Delta\phi_x/\Delta x$, and $k_z = \Delta\phi_z/\Delta z$, where k_x is the horizontal wavenumber in the X-direction and k_z is the vertical wavenumber in the Z-direction and $\Delta\phi_x$, $\Delta\phi_z$ are the autospectral phase differences in the respective X and Z directions. Although k_x and k_z computations do not correspond to the same location, they do suggest that those in the X-direction are of an order of magnitude less than those in the Z-direction. This is consistent with the laboratory set up of a vertically stratified water column in the wave tank, which inhibits propagation in the vertical. As a result, the spatially inhomogeneous turbulence undergoes a transition to two-dimensional motion under the action of stratification, and there would be a propensity for a stronger horizontal propagation, as suggested by the recent stratified shear flow experiments[17,18].

However, the basic observations which suggest that turbulence dominates in the near field of the source and

Table 2. Wavenumber k_z (values in rad/cm)

	Horizontal distance (cm)							
	10	30	50	70	10	30	50	70
	Frequency = 0.0244 Hz				Frequency = 0.1465 Hz			
E–F	2.001	0.107	0.517	–0.231	0.379	1.255	–1.937	–0.446
D–E	–1.785	1.531	0.187	–0.094	–0.624	–0.401	1.473	–0.427
C–D	1.606	–1.757	–1.944	0.427	1.292	0.259	–0.345	1.288
B–C	–0.004	1.303	1.795	1.412	–0.273	–1.091	–0.685	–1.917
A–B	0.221	–1.463	–0.490	0.251	–0.941	0.266	0.073	0.968

The origin of the horizontal distances is the source of forcing.
A–F denote vertical locations separated by 2.5 cm.

internal waves generated by the collapse of the turbulent wake dominate in the far field (see the raw time series plots in Figure 4 and low passed plots in Figure 5) are consistent with earlier studies[19,20]. Close to the centre of the disturbance, internal waves generated due to the displacement effects of the propeller propagate in a narrow region along the axis of the generating source[21]. This feature is reflected in the time series plots of run D.

The spectral observations in Figure 6 show that near the source (runs A & B), turbulence manifests (broad band spectra) and along the axis of the source (run D), internal wave signatures manifest as higher energy in the lower frequencies of the spectra, rapidly decaying towards higher frequencies. The spectral peaks associated with different frequencies at different locations are indicative of the multiple interactions between various components creating nonlinear energy transfers[22,23].

Conclusions

A novel optoelectronic technique to directly monitor internal waves in the laboratory was demonstrated. The technique is highly sensitive and is free from the conventional experimental constraints. It can be easily adapted for use in field experiments.

1. Garrett, C. J. R. and Munk, W., *Geophys. Fluid Dyn.*, 1975, **2**, 225–264.
2. Woods, J. D., *J. Fluid Mech.*, 1968, **32**, 791–800.
3. Wunsch, C., *J. Geophys. Res.*, 1975, **80**, 339–343.
4. Katz, E. J., *J. Phys. Oceanogr.*, 1973, **3**, 448–457.
5. Cairns, J. L., *J. Geophys. Res.*, 1975, **80**, 299–306.
6. Bell, T. H., *J. Geophys. Res.*, 1976, **81**, 3709–3714.
7. Müller, P., Olbers, D. J. and Willebrand, J., *J. Geophys. Res.*, 1978, **83**, 479–500.
8. Maxworthy, T. and Browand, F. K., *Annu. Rev. Fluid Mech.*, 1975, **7**, 273–305.
9. Kasischke, E. R., *IEEE J. Oceanic Engg.*, 1990, **15**, 126–130.
10. Lauterborn, W. and Vogel, A., *Annu. Rev. Fluid Mech.*, 1984, **16**, 223–244.
11. Sell, J. A., *Photo Thermal Investigations of Solids and Liquids*, Academic Press, London, 1989.
12. Rajasree, K., Vidyalal, V., Radhakrishnan, P., Nampoori, V. P. N. and Vallabhan, C. P. G., *Meas. Sci. Technol.*, 1992, **4**, 435–437.
13. Mandells, A., Takaue, R., Chen, Z., Szurmak, J. and Baines, W. D., *Anal. Sci.*, 1992, **8**, 131–136.
14. Mowbray, D. E. and Rarity, B. S., *J. Fluid Mech.*, 1976, **28**, 1–16.
15. Stevenson, T. N., *J. Fluid Mech.*, 1973, **60**, 759–767.
16. Holder, D. W. and North, R. J., *Schlieren Methods*, A National Physical Laboratory Publication, Middlesex, UK, 1963.
17. Lighthill, J., in *Waves in Fluids*, Cambridge University Press, London, 1978, p. 505.
18. Chu, V. H. and Baddour, R. E., *J. Fluid Mech.*, 1984, **138**, 353–379.
19. Schooley, A. H. and Steward, R. W., *J. Fluid Mech.*, 1963, **15**, 83.
20. Wu, J., *J. Fluid Mech.*, 1969, **35**, 531–544.
21. Sherman, F. S., Imberger, J. and Corcos, G. M., *Annu. Rev. Fluid Mech.*, 1978, **10**, 267–288.
22. Martin, S., Simmons, W. F. and Wunsch, C. I., *J. Fluid Mech.*, 1972, **53**, 17–44.
23. McEwan, A. D. and Robinson, R. M., *J. Fluid Mech.*, 1975, **67**, 667–687.

ACKNOWLEDGEMENTS. Assistance from T. P. Muralidharan, NPOL is gratefully acknowledged. KR and VV thank UGC and CSIR for their research fellowships. PR, VPNN and CPGV thank MHRD, Govt of India for financial grant. RT and PNA thank Director, NPOL for the encouragement and facilities provided.

Received 7 October 1994; revised accepted 26 September 1995.

Habitat suitability analysis of *Nemorhaedus goral* – A remote sensing and geographic information system approach

P. S. Roy, Shirish A. Ravan*, N. Rajadnya†, K. K. Das, Abhineet Jain and Sarnam Singh

Indian Institute of Remote Sensing, 4, Kalidas Road, P. B. 18, Dehradun 248 001, India
*World Wide Fund for Nature (India), 172-B, Lodi Estate, New Delhi 110 003, India
†Forest Department, Govt. of Maharashtra, Divisional Forest Office, Kolhapur, India

With the decreasing size of habitat and increasing fragmentation, it has become essential to develop species-specific habitat suitability maps. Such an approach for endangered species is a priority task for India's conservation programme. The mountain goat (goral) is an endangered species. The viable population of goral (*Nemorhaedus goral*) on Siwalik mountain is now confined to only Rajaji National Park, Uttar Pradesh. The goral habitat is restricted to specific land attributes, viz. landcover, slope, water availability and human habitation. Human settlements, biotic pressure (grazing and lopping) and erosional subsidence on hills have threatened goral habitat. The study is an attempt to model habitat suitability using spatial landscape patterns (interspersion and juxtaposition) and specific habitat requirements (restrictive factors). The approach has been tested in the part of Rajaji National Park for goral.

HABITAT is a place occupied by a specific population within a community of populations[1]. Habitat selection is an important part of an organism's life history patterns. The worldwide destruction of the natural environment is reducing the number of species and the amount of genetic variation within individual species. Major threats to biodiversity include habitat alterations, overharvesting, pollution, climate change, introduced species and human population increase. Of these, habitat alterations due to overuse, deforestation and consequent habitat loss, fragmentation and degradation are the primary factors[2]. There is a need to identify priority areas for conserving wildlife habitat. India's conservation effort is evolved around the UNESCO protected area concept by creating biosphere reserves, national parks, and sanctuaries. About 3.4% of India's geographical area is under protection for '*in situ*' biodiversity conservation. The approach takes into account endangered species and habitats. However, with mounting pressure on the protected areas and species, the approach is under review. It is now considered imperative to analyse baseline data on the physical set-up, vegetation and socioeconomic environment to ensure the effectiveness of the conservation effort.

Unfortunately, the study area (Rajaji National Park, Uttar Pradesh, India) has suffered from habitat loss, near fragmentation and degradation. The area around Rajaji National Park has had a twofold increase in human population from 1951. The increase in the human population has accentuated the biotic pressure on the forest[3].

The most characteristic Himalayan mammal of this park is the goral, a mountain goat. Johnsingh[3] has described the animal and its habitat requirements. The goral (*Nemorhaedus goral*) is a stocky goat antelope, 65–70 cm at the shoulder and 25–30 kg in weight. Both sexes have horns and a conspicuous white throat patch.

Goral habitat

Three species of goral inhabit the Himalaya, northeast part of India, Burma, China and Thailand[4]. The Himalayan goral (*Nemorhaedus goral*) is found from 200 msl in the Siwalik Hills to 4000 msl in the main Himalayan range[5]. The Siwalik forest landscapes of the Himalayas are exposed to various degradational processes due to their fragile geomorphology and heavy human dependency. Rajaji National Park is an attempt to protect the representative biogeographic unit (Gangetic plain). Goral habitat has been extensively studied and it is established that goral has a very specific habitat requirement[3,6–10]. The habitat size has been shrinking due to fast degradation caused by the lopping of trees, grazing and cutting of grass, and erosional subsidence of slopes. This has threatened the survival of goral.

The habitat parameters required for goral were summarized based on the field observations and literature:

– Goral is not seen near human habitation.

– Goral is a grazer. It selects vegetation with good grass cover.
– It avoids lesser slope angle categories and prefers areas with higher slope.
– Escape terrain is either a cliff, steep slope or '*nalah*'/gully followed by broken/rocky terrain. But cliffs are one of the most important escape terrains; forest cover seems to be the next highly used escape routes[8].
– Goral avoids areas with extensive shrub cover. This was related to forage availability as well as its anti-predatory strategy.

In spite of the species being endangered, no attempts have so far been made to analyse spatial habitat suitability and the available habitat size.

Habitat evaluation approach

The colonization pattern of the animals is related to the following parameters: cover, physical set-up and land use. The habitat suitability analysis based on these parameters is considered essential to protect the species. Since long time, ground methods have been used to evaluate the habitat using habitat suitability index based on ground observations[11,12]. However, these methods do not provide any spatial information. Aerospace remote sensing has been used effectively to characterize the habitat[13–15]. Subsequent attempts have also been made in India to evaluate wildlife habitats and related parameters using remote sensing data[16–18]. These studies, however are subjective and do not take into account spatial patterns and diversity of landscape. Researchers have used Geographic Information System (GIS) to handle multi-layered data and analyse spatial patterns statistically[19]. The model can be driven for different combinations of the controlling factors identified from a GIS database to predict suitability[20].

This study presents an approach to identify goral (*Nemorhaedus goral*) habitat based on analysis of spatial patterns of landscape. A remote-sensing-derived vegetation map, water sources, human settlement, road infrastructure and digital terrain model are the basic inputs for the study. The habitat suitability has been derived based on spatial patterns of landscape and the specific habitat requirements of goral.

Study area

The study area is located in Dhaulkhand range of the Rajaji National Park (77°55′35″E to 78°4′42″E longitude and 29°59′45″N to 30°8′48″N latitude) covering an area of 149.6 km^2 (Figure 1). The topography of the area varies with altitude, ranging from 320 m above msl in submontane area to above 800 m in the Siwalik ranges. The climate is subtropical. Rainy season commences from late June and continues up to the middle of September with the heaviest rainfall recorded in August (range 1200–1500 mm). The average annual rainfall is about 1270 mm. The temperature varies from 13.1°C (in January) to 38.9°C (in June). The vegetation mainly belongs to the northern tropical dry deciduous forest and sub-tropical forest[21]. Among fauna, herbivorous game animals of the deer family, cheetal (*Axis axis*), hog deer (*Axis pornicus*) are common. Large mammals like elephants are abundant.

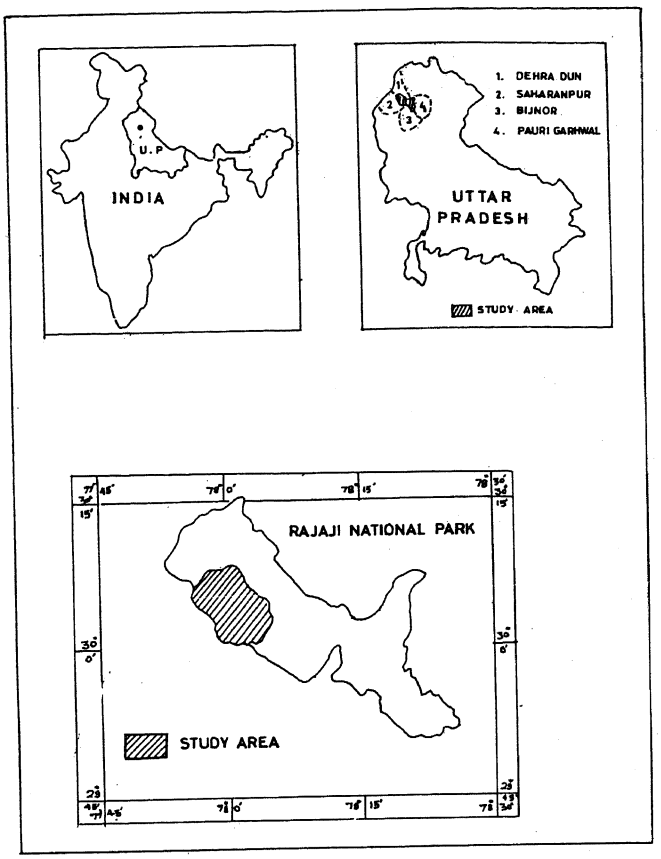


Figure 1. Location map of the study area.

Materials and methods

Input data

Indian Remote Sensing Satellite-1B Linear Imaging Self Scanning Sensors (IRS–1B LISS II) (with ground resolution of 36.25 m) false colour composite (FCC) of bands 4, 3 and 2 have been used to prepare the vegetation map on a 1:50,000 scale. The vegetation types and associated land use details interpreted are sal (*Shorea robusta*) (S), sal mix (SM), miscellaneous (M), miscellaneous with grass (MG), chir (*Pinus roxburghii*) (C), chir with grass (CG), scrub (Sc), riverine (r), plantation

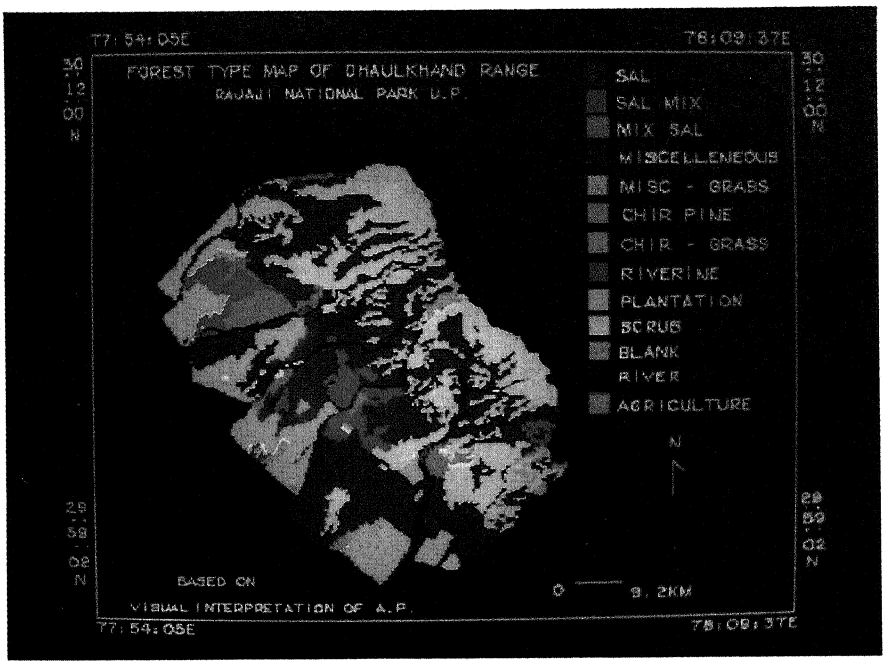


Figure 2. Forest type map of Dhaulkhand Forest Range, Rajaji National Park, UP.

Table 1. Locations where gorals were sighted either feeding/resting

	Habitat parameters			
Location	Slope (degrees)	Forest type	Water	Habitation
Dhaulkhand ridge near Sayyawala point	35–60	Misc. with grass, Misc.	Within 1 km	Nil
Dhaulkhand ridge about 200 m NW from point 1	35–60	Misc. with grass, Misc.	Within 1 km.	Nil
Ridge between Gularia and Tangsot towards Tangsot	>70	Misc. with grass	Within 100 m	Nil
Gulari ridge	>60	Misc.	Within 100 m	Nil
Ridge between Gulari and Tangsot point towards Dhaulkhand road	>50	Misc. with grass	Within 100 m	Nil

(Pt), agriculture (Ag), drainage and water bodies (Figure 2). The terrain features like contours and roads were obtained from Survery of India toposheets. The known habitat parameters for goral (from field and literature) were used along with the spatial information in GIS.

GIS approach

PAMAP GIS configured on PC-based platform has been used in the present study. The maps showing vegetation/land use, contours, water and human habitation were digitized. The contours were interpolated to calculate digital terrain model (DTM) and the slope map was derived. Finally, the spatial information along with restrictive factors (Table 1) were modelled using overlay and neighbourhood techniques. The study approach is presented in Figure 3.

Results

Spatial pattern analysis

The vegetation landscape is analysed to determine habitat interspersion and juxtaposition. The interspersion (Is) is a measure of the spatial intermixing of habitats/land use and is calculated in a nonspecies-specific

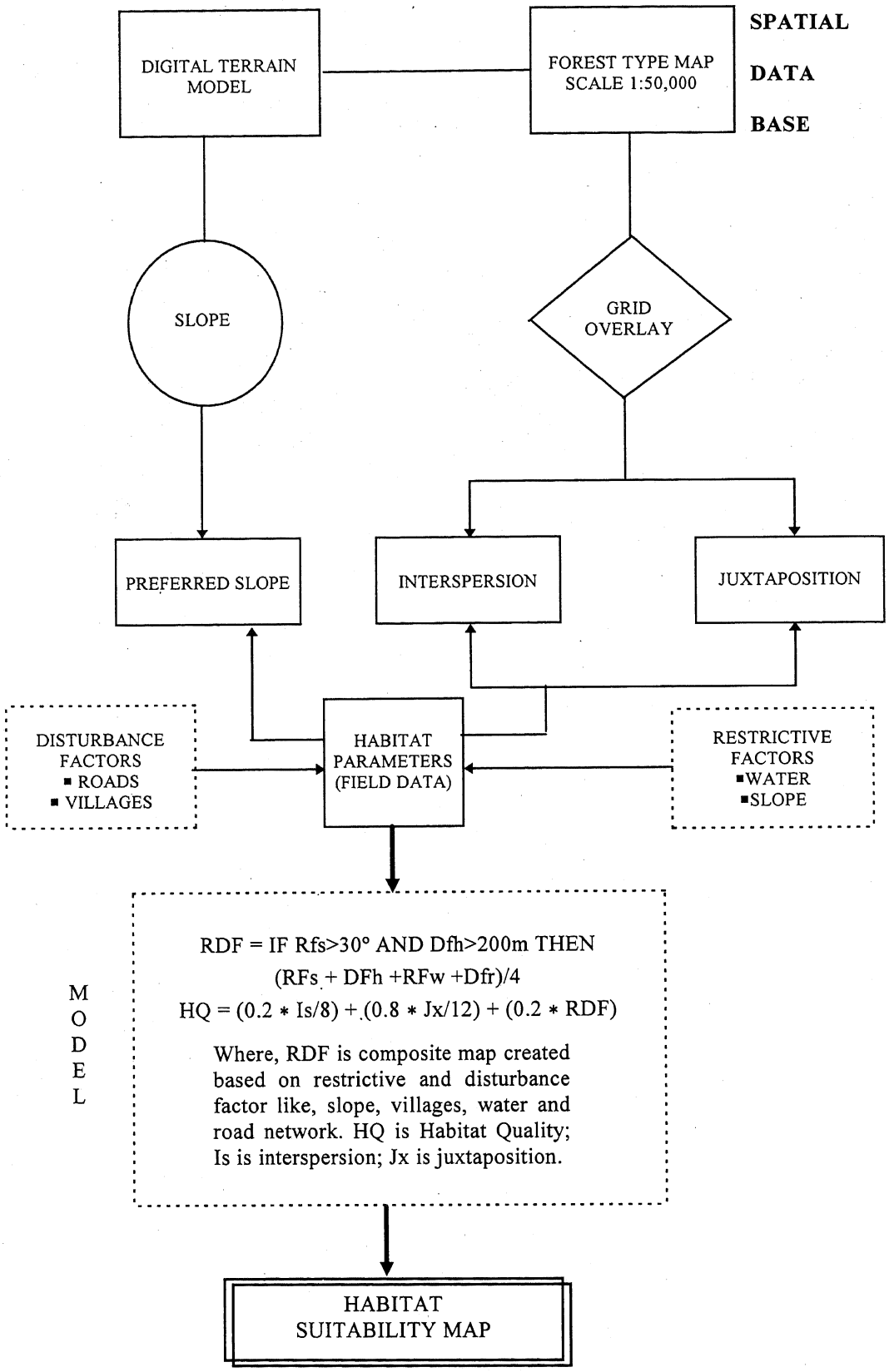


Figure 3. Approach for habitat suitability analysis.

manner[12,22]. The juxtaposition (Jx) is a measure of the proximity of habitat types and relative importance of adjacency[12,23]. Thus, Jx is a measure species-specific to goral. This was accomplished by defining a grid which was placed on a forest type map. Based on the field observation on the habitat size, grids of 20 s × 20 s were used for the present study. Grid sizes can be changed as per the required level of details. Gridwise interspersion and juxtaposition were calculated using a software developed in database management system (dBase III+) interfaced with PAMAP GIS. The weightage factors were assigned to different edges based on the order of preference by goral.

Measure of interspersion

A window of size 3 × 3 grids panned over vegetation map and an interspersion value was assigned to the central grid to prepare the interspersion map. Interspersion of central cells is calculated as the number of surrounding grid cells that differ from the central cell. The interspersion of a given window is 6. The maximum limit of interspersion is 8.

MG	MG	M
S	MG	M
S	Sc	S

Measure of juxtaposition

The juxtaposition of the central grid cell is calculated by comparing the class of the central cell with an adjacent cell as per the weighted preference of the species on use of the habitat. The length of the edge (diagonal or horizontal/vertical) is also provided as a weightage. The habitat preference and weightage criteria are given as follows.

Goral habitat preference of vegetation cover in descending order is:
Miscellaneous forest with grasses on slopes (MG) – Chir-Pine with grasses on slopes (CG) – Miscellaneous (M) – Chir-Pine (C) – Miscellaneous sal (MS) – Sal miscellaneous (SM) – Scrub (Sc) – Blanks (B) – Riverine (R) – Sal (S) – Plantation (Pt) – Agriculture (Ag). The weightages assigned to different edges are given in Table 2. Calculation

Table 2. Weightage assigned to different edges of vegetation types

Vegetation type factor	Association	Weightage
Miscellaneous forest with grasses on slopes	Any cover	0.9
Chir-pine with grasses on slopes	Any cover	0.9
Miscellaneous forest	Chir-pine with grasses, miscellaneous, scrub and blanks	0.9
Miscellaneous forest	Any other cover	0.6
Miscellaneous sal	Any cover	0.6
All other combinations		0.2

Table 3. Measure of juxtaposition

Edge combination	Weightage factor (specific to goral)	Quantity	Juxtaposition (Jx)
M/Sc	0.6	3	1.8
MG/Sc	0.9	3	2.7
C/Sc	0.2	2	0.4
R/Sc	0.2	1	0.2
Sc/Pt	0.2	2	0.4
Sc/Ag	0.2	1	0.2

Note: Diagonal count as 1, while vertical or horizontal count as 2.

of juxtaposition for the central grid of the given window is given in Table 3 and is equal to 5.7.

M	M	MG
S	Sc	C
Ag	Pt	R

The juxtaposition in the study area ranges from 2.4 to 10.8. The maximum limit of juxtaposition is 12.

Restrictive (RF) and disturbance (DF) factors

The factors that are essential for wildlife species are called restrictive factors. The presence of this factor is indicated by value one and has no impact on habitat suitability index; otherwise the factor is zero and, consequently, habitat suitability index becomes zero. Intermediate values can also be assigned based on field conditions and habitat requirements of wildlife. Based on habitat use by goral, four restrictive factors were determined.

Water availability (RFw). The availability of water within 1 km is the most important requirement of goral habitat. This was accomplished by generating the map showing a distance buffer around water sources and assigning ratings (0 to 1) based on the proximity.

Slope (RFs). Ratings (0 to 1) were given for slope classes. The rating values are 0, 0.8 and 1 for slope classes <30°, 30°–50° and >50°/cliffs, respectively.

Remoteness from human habitation (DFh). The human habitation is considered a disturbance factor. The presence of villages may alter the suitability of the grid. This was accomplished by generating the map showing a distance buffer around village locations and assigning ratings (0 to 1) based on proximity.

Remoteness from roads (DFr). The roads, although do not have significant impact on habitat, are considered a potential habitat disturbance factor. This was also accomplished by generating the map showing a distance buffer around roads and assigning ratings (0 to 1) based on proximity.

The cost surface developed using these four factors is called restrictive and disturbance factor map (***RDF***). This map was generated using the logic given below:

CREAT **RDF** IF RFs >30° and IF DFh > 200 m THEN (RFs + DFh + RFw + DFr)/4

Figure 4 displays the quality of habitat based on slope and proximity to water, road and human habitation. This map was used along with interspersion and juxtaposition maps to determine the final habitat suitability map.

Habitat suitability index

The habitat suitability is a function of interspersion, juxtaposition and restrictive factors. It is determined based on different weightages assigned to each parameter. The equation used for calculating the habitat suitability index is given below.

habitat index = (0.2*Is/8) + (0.6*Jx/12) + (0.2*RDF).

Thus, each cell was assigned a habitat suitability index ranging from 0 to 1. Histogram equalization was followed to cluster showing different habitat quality zones, viz. highly suitable, moderately suitable, less suitable and not suitable (Figure 5).

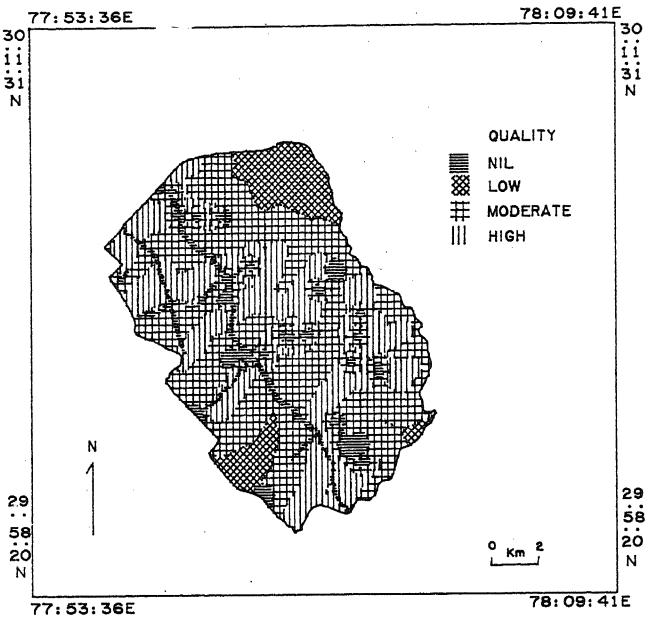


Figure 4. Cost surface created using restrictive and disturbance factors.

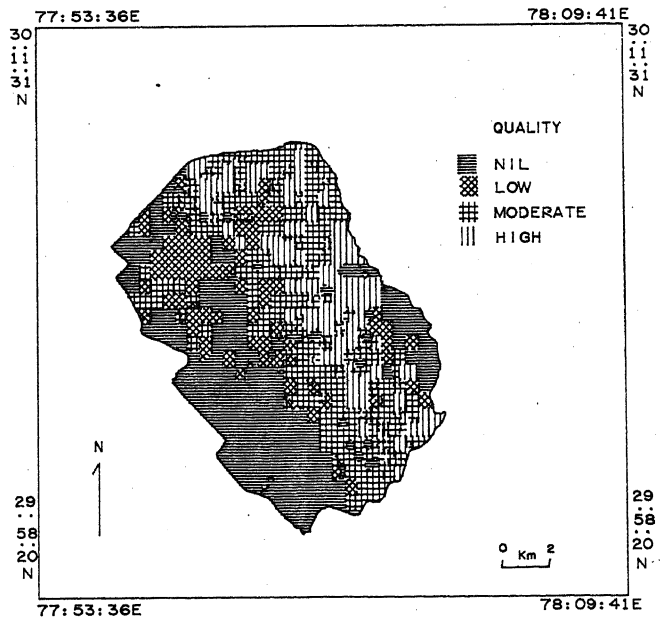


Figure 5. Habitat suitability map for goral in Dhaulkhand range, Rajaji National Park, UP.

Model comparison with field data

The suitable locations for goral noted down during the field work were superimposed on the habitat suitability map. Out of 10 locations, 5 are in most suitable areas, 3 in moderately suitable areas, 2 in less suitable areas and none in unsuitable areas.

The area under each category in percentage is as follows:

Habitat quality	Percentage area
Most suitable	35.22
Moderately suitable	15.15
Low suitable	29.47
Not suitable	20.16.

Discussion

With decreasing habitat and increasing fragmentation, it has become essential to develop species-specific habitat suitability maps. Such an approach for endangered species is a priority task for India's conservation programme. The mountain goat is an endangered species. The viable population of goral on Siwalik mountain is now confined only to Rajaji National Park. In the recent years, increasing human impact within the national park has threatened the very existence of goral in Siwalik. The present study is an attempt to develop an approach for habitat suitability mapping using spatial landscape patterns and specific habitat requirements. The model uses goral habitat parameters reported by Green[8] and Johnsingh[3].

The habitat studies indicate that homogeneity/heterogeneity of landscape, vegetation types, their association and edges characteristics are important parameters. Abundance of these species may be considered a consequence of edges where types of food and cover come together. The edge effect may be defined as the tendency for an increase in variety and density of an organism at a community junction. The present study tries to model it through measurement of interspersion and juxtaposition. The weightages of these two are based on species habitat preference. The conceptual framework on the use of these parameters has been discussed by Mead *et al.*[22] and Lyon[23]. Since the animal is shy to human presence, human habitation and road

proximity are also used in the model as parameters. The topographical habitat features and the availability of water have also been used as restrictive factors in the present model. The present study attempts to implement habitat suitability approach for the mountain goat.

Conclusion

India has large biodiversity. However, recent population growth trends have put tremendous pressure on the habitats. The situation calls for urgent need to develop comprehensive database of the protected areas, which are temples of *in situ* conservational efforts. It is necessary to evaluate the available habitat size, quality and socioeconomic constraints. The present approach can provide information about the suitability of endangered species having very specific habitat requirements. There is need to implement such an approach for the species which are dwindling and becoming endangered due to shrinking habitat size.

1. Smith, R. L., *Ecology and Field Biology*, Harper and Row Publishers, New York, 1974.
2. McNeely, J. A., Miller, K. R., Ried, W. V., Mittermeir, R. A. and Werner, T. B., Conserving World's Biological Diversity, Washington, DC, IUCN/WRI CI/WWF-US/World Bank, 1990.
3. Johnsingh, A. J. T., *Sanctuary*, 1991, **XI**, 14–25.
4. Groves, C. P. and Grubb, P., in *The Biology and Management of Mountain Ungulates* (ed. Lovari, S.), Croom Helm, 1985, pp. 45–50.
5. Schaller, G. B., *Mountain Monarchs*, University of Chicago Press, Chicago, 1977.
6. Prater, S. H., *The Book of Indian Animals*, Bombay Natural History Society, Oxford University Press, 1980.
7. Lovari, S., *Int. J. Zool.*, 1986, 53.
8. Green, M. J. B., *J. Zool. London*, 1987, **B1**, 693–719.
9. Cavallini, P., *J. Bombay Nat. Hist. Soc.*, 1992, **89**, 302–307.
10. Mishra, C., Habitat use by goral (*Nemorhaedus goral*) in Majhatal Harsang Wildlife Sanctuary, Himachal Pradesh, India. M Sc Dissertation, Wildlife Institute of India, Dehradun, 1993.
11. Balda, R. P., Vegetation structure and breeding bird diversity, For. Res. Tech. Report No. 1, USDA, Arizona, 1975.
12. Giles, R. H., *Wildlife Management*, Freeman and Co., San Francisco, 1978.
13. William, D. L. and Nelson, R. F., *IEEE Trans. Geosci. Remote Sensing*, 1986, **GE-24**, 130–138.
14. Porwal, M. C. and Roy, P. S., *Environ. Conserv.*, 1991, **18**, 45–50.
15. Ravan, S. A. and Roy, P. S., *Curr. Sci.*, 1994, **64**, 309–315.
16. Kushwaha, S. P. S. and Unni, N. V. M., Application of remote sensing technique in forest cover monitoring and habitat evaluation – A case study in Kaziranga National Park, Assam, Proc. Seminar-cum-Workshop on Wildlife Habitat Evaluation Using Remote Sensing Technique (Dehradun), 1986, pp. 238–247.
17. Parihar, J. S., Kotwal, K. C., Panigrahy, S. and Chaturvedi, N., Study of wildlife habitat using high-resolution space photographs – A case study of Kanha National Park, ISRO-SP-17-86, Results from the Joint Indo-Soviet Remote Sensing Experiment, TERRA on board Salyut-7, 1986.
18. Porwal, M. C. and Roy, P. S., Wildlife habitat analysis in Kanha National Park using remote sensing, Presented in the National Symposium on Remote Sensing of Environment (ISRS), Madras, 1991.
19. Akçakaya, *GIS World*, 1994, November, 36–49.
20. Baker, W. L., *Landscape Ecol.*, 1989, **2**, 111–133.
21. Champion, H. G. and Seth, S. K., *A Revised Survey of the Forest Types of India*, Manager of Publications, Govt. of India, New Delhi, 1968.
22. Mead, A. R., Sharik, T. L., Prisley, S. P. and Heinen, J. T., A computerised spatial analysis system for assessing wildlife habitat from vegetation maps, 47th Annual Meeting of the American Society of Photogrammetry, Washington, 1981.
23. Lyon, J. G., *Photogramm. Engg. Remote Sensing*, 1983, **49**, 245–250.

ACKNOWLEDGEMENTS. We are grateful to Prof. B. L. Deekshatulu, Director, National Remote Sensing Agency, Hyderabad and Prof. S. K. Bhan, Dean, Indian Institute of Remote Sensing, Dehradun for providing all the facilities required for this study. We are also grateful to Dr A. J. T. Johnsingh, Head, Wildlife Biology Faculty, Wildlife Institute of India, Dehradun for providing valuable suggestions and review.

Received 3 July 1995; revised accepted 22 September 1995

Carbon and oxygen isotopic records of planktonic and benthic foraminifera from a new deep-sea core of the northeast Indian Ocean

S. Masood Ahmad

National Geophysical Research Institute, Hyderabad 500 007, India

Stable carbon and oxygen isotopic records of the planktonic foraminifera (*Globigerinoides ruber*), covering the past ~100,000 years, have been obtained from a new deep-sea core of the Bay of Bengal. High glacial-to-interglacial $\delta^{18}O$ amplitude (2.0‰) in the surface waters of this region provide evidence for the increased salinity (by ~2‰) in the glacial surface waters of this region[1,2]. In contrast, the surface water $\delta^{13}C_{TIC}$ did not change significantly during the glacial-to-interglacial transition.

Examination of the isotopic record of benthic foraminifera (*Cibicidoides* and *Uvigerina*) shows higher (1.7‰) glacial-to-interglacial $\delta^{18}O$ amplitude. This is attributed to the decrease in deep-water temperature arising from changes in source characteristics during the LGM. The glacial-to-interglacial shift in $\delta^{13}C_{TIC}$ (~0.3‰) in the deep Bay of Bengal is close to the global

$\delta^{13}C$ change, in conflict with the previous reports that suggested marked enrichment of nutrients in deep waters of the northeast Indian Ocean.

STABLE carbon and oxygen isotopic analyses in the planktonic and benthic foraminifera from the deep sea cores of the northern Indian Ocean have shown significant changes in the chemical characteristics of the surface and deep-water masses since the last glacial maximum (LGM)[1–7]. Previously published oxygen isotopic records of the planktonic foraminifera from the Northeast Indian Ocean (NEIO) are suggestive of much higher salinity as compared to today in the glacial surface waters of this region due to a reduction in riverine input[1,7]. A large influx of pollen from the Irano–Turanian Steppe and high percentage of quartz observed in the glacial sections of sediment cores also support prevalence of high aridity in the region and a weakening of the southwest monsoon during the last glaciation[8,9].

Heavy precipitation during the southwest monsoon over the Indian subcontinent is collected by the major rivers (Ganges, Brahmaputra, Salween and Irrawady) which drain into the northeast Indian Ocean. As a result, a strong north–south salinity gradient forms in the Bay of Bengal, which is reflected in the $\delta^{18}O$ distribution pattern of the Holocene planktonic foraminifera[1]. The deep Bay of Bengal is filled with water originating in the circumpolar region and transported northward by the two western boundary currents[10]. The water found at the site of the core SK-72 (water depth 3246 m) has an important component from the relatively warm North Atlantic Deep Water (NADW).

The hydrographic structure of the deep and intermediate water in the northern Indian Ocean experienced considerable changes during the last glacial maximum[3,5]. The $^{13}C/^{12}C$ and $^{18}O/^{16}O$ ratios in shells of benthic foraminifera showed that, during the last glaciation, water column in the northern Indian Ocean was highly stratified, with a well-marked boundary between the intermediate- and deep-water masses[3], in contrast to the present conditions, when the $\delta^{13}C$ gradient is very small (0.1‰)[11]. To explain this glacial stratification, a mechanism involving the transfer of nutrients from the intermediate to deep waters was suggested[12,13]. But more recent data from intermediate-depth cores of the NEIO showed that the glacial $\delta^{13}C$ gradient was not as high as reported earlier[5,6]. The temperature of glacial intermediate- and deep-water masses in NEIO was also lower by at least 1–1.5°C compared to the present bottom water of this region[3,5,14].

The distribution of carbonates in deep-sea sediments is controlled by the depth of the lysocline, where the rate of calcite dissolution increases sharply and by the carbonate compensation depth (CCD), below which calcium carbonate does not occur. The biological productivity, water depth, chemistry and dilution by non-carbonate components are the main factors influencing

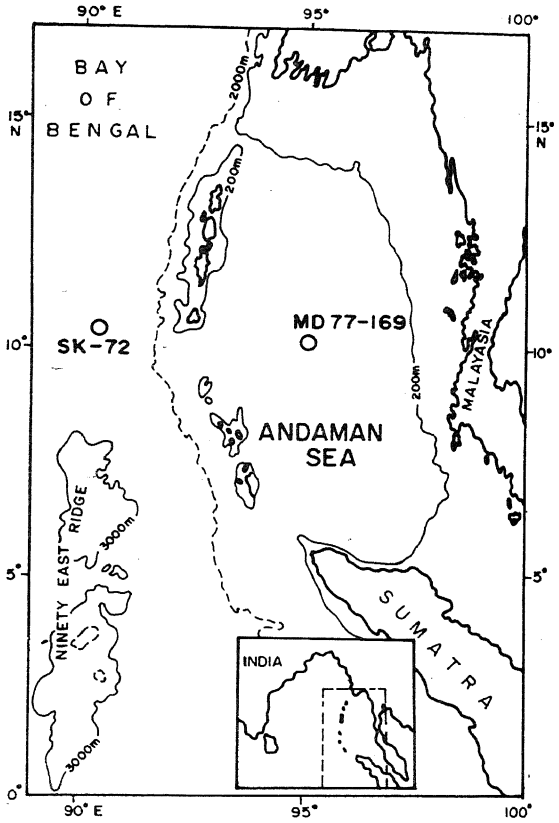


Figure 1. Location map of the cores SK-72 and MD 77-169.

carbonate content of the sediments[15]. In addition, the deep-water circulation pattern also plays an important role in determining carbonate distribution[15,16]. As a result of these changes, the depth of the lysocline varies considerably, both locally as well as globally.

The box core SK-72 was retrieved from the Bay of Bengal (10°43′N and 90°27′E) during February/March 1992 at a water depth of 3246 m (Figure 1). The core was subsampled for this study at 10–20 cm intervals except within the Holocene section, where samples were taken at 5 m intervals. The sediment samples were soaked in distilled water overnight and subsequently washed through a 150 μm sieve and dried at 60°C. About 15–20 foraminiferal tests in the size range of 250–315 μm were then picked under a microscope for isotopic measurements.

The separated specimens were ultrasonically cleaned to remove adhering particles and roasted under vacuum at 400°C for 1 h. The CO_2 was extracted by the reaction of foraminiferal calcite with 100% phosphoric acid at 50°C on an on-line system and the isotopic ratios were measured using a VG Micromass 903 mass spectrometer. The precision based on the replicate measurements of the internal laboratory standard MMB for both carbon and oxygen was better than ±0.1‰. The results are expressed relative to PDB standard (Table 1). *Cibicidoides* oxygen isotopic values were converted into *Uvigerina* scale by adding 0.64‰ (ref. 17), while carbon isotopes

Table 1. Carbon and oxygen isotopic composition of planktonic (*G. ruber*) and benthic (*Cibicidoides* and *Uvigerina* spp.) foraminifera for core SK-72. Isotopic values are given in per mil against PDB standard

Depth (cm)	Spp.	$\delta^{13}C$	$\delta^{18}O$	Depth (cm)	Spp.	$\delta^{13}C$	$\delta^{18}O$
0	*Ruber*	0.79	−2.52	0	*Cibicidoides*	0.36	2.67
5	*Ruber*	0.87	−2.53	5	*Cibicidoides*	–	2.42
20	*Ruber*	1.21	−2.23	10	*Cibicidoides*	0.29	2.41
60	*Ruber*	0.97	−0.63	20	*Cibicidoides*	0.48	3.11
80	*Ruber*	1.15	−0.37	30	*Cibicidoides*	0.12	3.58
120	*Ruber*	1.14	−0.91	35	*Cibicidoides*	0.16	3.68
140	*Ruber*	1.34	−0.62	80	*Uvigerina*	−0.69	5.03
160	*Ruber*	1.21	−0.67	100	*Uvigerina*	−0.80	5.04
180	*Ruber*	1.03	−0.96				
200	*Ruber*	1.29	−1.11				
220	*Ruber*	1.13	−1.11				
240	*Ruber*	1.81	−0.78				
260	*Ruber*	1.13	−0.63				
280	*Ruber*	1.08	−1.03				
300	*Ruber*	1.23	–				
320	*Ruber*	1.26	−0.92				
340	*Ruber*	1.48	−1.31				
360	*Ruber*	1.42	−1.08				
380	*Ruber*	1.22	−1.55				
400	*Ruber*	1.15	−1.39				

Table 2. Age against depth data for core SK-72. Ages are estimated from the time scale derived for the normalized SPECMAP curve by Imbrie *et al.*[19]

Event	Sub-bottom depth (m)	Age 10^3 yr
1.1	0.2	6
2.0	0.4	12
2.2	0.8	18
3.1	1.2	24
3.3	2.0	53
4.0	2.39	59
4.2	3.0	65
5.1	3.4	80
5.2	3.6	87
5.3	3.8	94

of *Uvigerina* were brought to the *Cibicidoides* scale by adding 0.83‰ (ref. 18). Carbonate contents were determined by weight loss after leaching the samples with dilute (50%) hydrochloric acid. The reproducibility of the carbonate abundance estimate is better than ±2% based on six replicate measurements.

Age assignments and correlation of cores are based on the planktonic oxygen isotopic stratigraphy by using graphic correlation of the SPECMAP $\delta^{18}O$ record (Table 2)[19]. The average sedimentation rate in SK-72 is about 4.1 cm/ka. The planktonic record of SK-72 correlates well with another core (MD 77-169) of the northeast Indian Ocean[7].

The results of the oxygen isotopic analyses of the surface-dwelling planktonic foraminifera *Globigerinoides ruber* in SK-72 are plotted against time in Figure 2. The observed glacial-to-interglacial $\delta^{18}O$ amplitude exceeds the ice volume effect of ~1.2‰ (ref. 20) by 0.8‰. This could arise from either a cooling of surface waters by about 3°C (ref. 21) or from an increase in salinity by ~2‰ (ref. 22), or from a combination of both the factors. However, earlier studies have shown that the glacial surface temperature in the Bay of Bengal did not differ by more than 1°C from the present SST[1,22]. Therefore, the glacial enrichment in $\delta^{18}O$ could be attributed mainly to the increased salinity. Comparison of the $\delta^{18}O$ records between SK-72 and the Andaman Sea core MD77-169 indicates generally lower ^{18}O in the Andaman Sea, because core MD77-169 comes from a region which is greatly influenced by runoff from the rivers Irrawady and Salween.

The LGM to core top change in $\delta^{18}O$ for the benthic foraminifera (~1.7‰; Figure 3) is also considerably higher than the ice volume effect[20]. Since there is no evidence for any local deep water formation in NEIO during the last glaciation, the higher glacial–Holocene shift in benthic $\delta^{18}O$ in this region should be due to a decrease in deep-water temperature as a result of the changes in source waters[14,23]. It is quite likely that the glacial deep water in the Bay of Bengal had a greater contribution from the Antarctic Bottom Water (AABW) at the expense of warmer NADW as production of the latter was greatly reduced during the glacial times. The additional $\delta^{18}O$ enrichment (~0.5‰) in SK-72 implies a cooling by about 2°C; this is somewhat higher than the previous estimate (1.5°C)[3]. However, it must be noted that these computations assume no local salinity change in excess of that included in the global ice volume effect. If there was a local increase in salinity as well (which is likely given the dramatic increase in surface salinity during the LGM and the consequent downward diffusion of salt through the water column), the magnitude of cooling would be smaller.

The present geographic distribution of $\delta^{13}C$ in the deep sea is related to the oxygen and nutrient contents

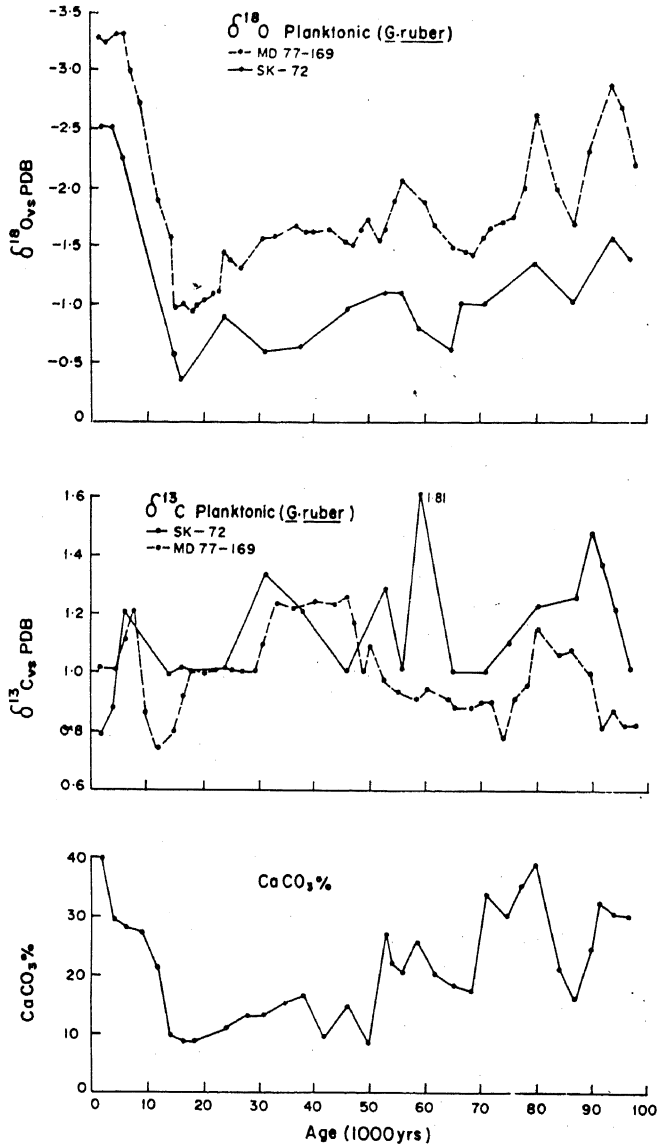


Figure 2. Carbon and oxygen isotopes in planktonic foraminifera (*Globigerinoides ruber*) and $CaCO_3$ percent in the cores SK-72 and MD 77-169 plotted against age.

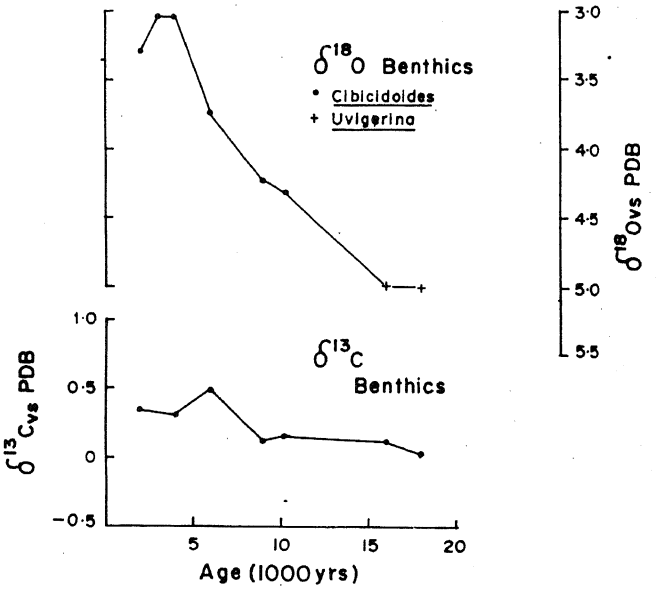


Figure 3. Stable isotopic record of the benthic foraminiferal genus *Cibicidoides* and *Uvigerina*. Oxygen isotopes of *Cibicidoides* corrected by adding 0.64‰ and carbon isotopes of *Uvigerina* corrected by adding 0.83‰.

and circulation pattern of the water masses[11]. However, during the last glaciation the deep-water circulation probably experienced considerable changes because of the decreased production of NADW[24,25]. Carbon isotopic data on benthic foraminifera belonging to the genera *Cibicidoides* and *Uvigerina* have shown that during the LGM deep-ocean $\delta^{13}C$ was lower by 0.32–0.46‰ compared to the late Holocene[4,18]. In SK-72 the maximum $\delta^{13}C$ difference (0.45‰) was observed between 20 and 100 cm levels. However, the mean glacial–interglacial $\delta^{13}C$ shift is only ~0.3‰ (Figure 3), which is very close to the global deep-water $\delta^{13}C$ shift. Recent studies from the intermediate-depth water cores from the North Indian Ocean have clearly demonstrated that the glacial mid-depth water masses in NEIO are also similarly depleted in ^{13}C by 0.3–0.4‰, compared to the Holocene[5,6]. Hence, the $\delta^{13}C$ gradient between the intermediate and deep waters was very similar to that observed today. It can, therefore, be concluded that the glacial NEIO did not experience a dramatic change in the vertical chemical structure (acute oxygen depletion and nutrient enrichment in the deep waters) as reported previously[3].

The carbonate content of SK-72 varies from ~8% to ~40% by weight (Figure 2). The carbonate content is high (39.4%) in the top portion of the core (0–5 cm), decreases rapidly to 8.5% at 100–105 cm and remains low down to 200 cm depth. Thereafter, the variation was relatively small. High carbonate contents correspond to interglacials (oxygen isotope stages 1 and 5), whereas low values are seen during the glacial stages 2 and 4. The carbonate content during isotopic stage 3 is also low. This $CaCO_3$ abundance pattern is consistent with the carbonate data of other cores in the North Indian Ocean[26,27]. The low $CaCO_3$ abundance during the glacial times could be attributed to the decrease in lysocline as well as dilution by terrigenous matter. As stated earlier, the glacial deep waters probably received a greater contribution of cooler AABW. The decrease in sea water temperature favours carbonate dissolution and this would have contributed to the observed decrease in sedimentary $CaCO_3$ content during glacial times.

1. Duplessy, J. C., *Nature*, 1982, **295**, 494–498.
2. Sarkar, A., Ramesh, R., Bhattacharya, S. K. and Rajagopalan, G., *Nature*, 1990, **343**, 549–551.
3. Kallel, N., Labeyrie, L. D., Juillet-leclere, A. and Duplessy, J. C., *Nature*, 1988, **333**, 651–655.
4. Curry, W. B., Duplessy, J. C., Labeyrie, L. D. and Shackleton, N. J., *Paleoceanography*, 1988, **3**, 317–341.

5. Naqvi, S. W. A., Charles, C. D. and Fairbanks, R. G., *Earth Planet. Sci. Lett.*, 1994, **121**, 99–110.
6. Ahmad, S. M. and Labeyrie, L. D., *Geo-Marine Lett.*, 1994, **14**, 36–40.
7. Fontugne, M. R. and Duplessy, J. C., *Palaeogeogr. Palaeoclimatol. Palaeoecol.*, 1986, **56**, 69–88.
8. Kolla, V. and Biscaye, P. E., *J. Sed. Petrol.*, 1977, **47**, 642–649.
9. Van Campo, E., Duplessy, J. C. and Rossignol-Strick, M., *Nature*, 1982, **296**, 56–59.
10. Warren, B. A., *Deep Sea Res.*, 1981, **28**, 759–788.
11. Kroopnick, P. M., *Deep Sea Res.*, 1985, **32**, 57–84.
12. Boyle, E. A., *Nature*, 1988, **331**, 55–56.
13. Boyle, E. A., *Annu. Rev. Earth Planet. Sci. Lett.*, 1992, **20**, 245–288.
14. Sarkar, A. and Bhattacharya, S. K., in *Oceanography of Indian Ocean* (ed. Desai, B. N.), Oxford and IBH, New Delhi, pp. 455.
15. Berger, W. H. and Killingley, J. H., *Science*, 1977, **197**, 563–566.
16. Farrel, J. W. and Prell, W. L., *Paleoceanography*, 1988, **4**, 447–466.
17. Shackleton, N. J. and Opdyke, N. D., *Quat. Res.*, 1973, **3**, 39–55.
18. Duplessy, J. C., Shackleton, N. J., Mathews, R. K., Prell, W. L., Ruddiman, W. F., Caralp, M. and Hendy, C. H., *Quat. Res.*, 1984, **21**, 225–243.
19. Imbrie, J., Hays, T. D., Martinson, D. G., McIntyre, A., Mix, A. C., Morley, J. J., Pislas, N. G., Prell, W. and Shackleton, N. J., in *Milankovitch and Climate* (eds Berger *et al.*), Hingam, Massachusettz, 1984, pp. 305.
20. Labeyrie, L. D., Duplessy, J. C. and Blank, D. L., *Nature*, 1987, **347**, 477–482.
21. Shackleton, N. J., *Nature*, 1967, **215**, 15–17.
22. CLIMAP Project Members, Map and Chart Series MC-36, Geological Society America, Boulder, CO, 1981.
23. Chen, J. J. and Prell, W. L., *EOS*, 1990, **71**, 1381.
24. Oppo, D. W. and Fairbanks, R. G., *Earth Planet. Sci. Lett.*, 1987, **86**, 1–15.
25. Duplessy, J. C., Shackleton, N. J., Fairbanks, R. G., Labeyrie, L. D., Oppo, D. and Kallel, N., *Paleoceanography*, 1988, **3**, 343–360.
26. Murray, D. W. and Prell, W. L., *Proc. ODP*, 1987, **117**, 343–363.
27. Divakara Naidu, P., *Palaeogeogr. Palaeoclimatol. Palaeoecol.*, 1991, **86**, 255–263.

ACKNOWLEDGEMENTS. I express my sincere thanks to R. Ramesh for his assistance in the isotopic measurements at P.R.L., Ahmedabad, and K. Gopalan for his encouragement and support to carry out this work. The support and help provided by P. K. Banerjee in obtaining the core samples is thankfully acknowledged. I thank the referees of *Current Science* for their valuable comments. Thanks are also due to S. Ramaswamy for providing me the samples of the core SK-72.

Received 15 February 1995; revised accepted 15 September 1995

Resistance to *Bacillus sphaericus* in *Culex quinquefasciatus* Say 1823

T. Adak, P. K. Mittal, K. Raghavendra, Sarala K. Subbarao, M. A. Ansari and V. P. Sharma

Malaria Research Centre, Indian Council of Medical Research, 22, Sham Nath Marg, Delhi 110 054, India

Biological control of *Culex quinquefasciatus* using *Bacillus sphaericus* was considered a practical solution because of its specific and prolonged killing action against mosquito larvae. To study the feasibility of *B. sphaericus* ('Spherix') in mosquito control, multicentric trials were undertaken. Initially, *B. sphaericus* was very effective but within a year, after 20–25 rounds of application, field populations of *Cx. quinquefasciatus* developed resistance up to 150-fold. Genetic studies revealed that resistance was recessive, autosomal and controlled by more than one gene. This is the first report on nature and mode of inheritance of resistance against *B. sphaericus* in mosquitoes.

CULEX quinquefasciatus Say, 1823, is the major vector in the transmission of lymphatic filariasis caused by the nematode *Wuchereria bancrofti* in the tropics. To control *Cx. quinquefasciatus*, *Bacillus sphaericus* is emerging as a promising larvicide[1,2]. Extensive multicentric field trials carried out by Malaria Research Centre in India showed that the spraying of 'Spherix' (a commercial preparation of *B. sphaericus* Russian strain B-101, Serotype H5a 5b;) at the rate of 1 g/m^2 provides control of *Cx. quinquefasciatus* breeding for 2–4 weeks[3,4]. After 20–25 rounds of continuous field application of *B. sphaericus* within a year, reduction in its efficacy against *Cx. quinquefasciatus* was noticed. In this communication we report the development of resistance against *B. sphaericus* in field populations of *Cx. quinquefasciatus* and its mode of inheritance in this species.

Adult mosquitoes from *B. sphaericus* sprayed and unsprayed areas of Farrukhabad and Ghaziabad (Uttar Pradesh), Madras (Tamil Nadu) and Panaji (Goa) were collected and transported to MRC laboratories in Delhi for colonization[5] and further studies. In addition to these field populations, two laboratory strains maintained since 1988, one each from Delhi state and Sonepat (Haryana) were also used in bioassay and genetic crosses. Bioassays were performed following the method of WHO[1] with necessary modifications[6]. One per cent stock suspension of Spherix in dechlorinated water was prepared fresh and was used immediately for the bioassays. Twenty-five late III/early IV instar larvae were placed in 250 ml of water containing the required dose of *B. sphaericus* for 24 h and scored for mortality. All the tests were carried out in two replicates at 27± 1°C. Dose–mortality responses of treated lines were analysed using probit regression method of Finney[7].

Dose–mortality regression of the field populations and two laboratory strains are given in Figure 1 and Table 1.

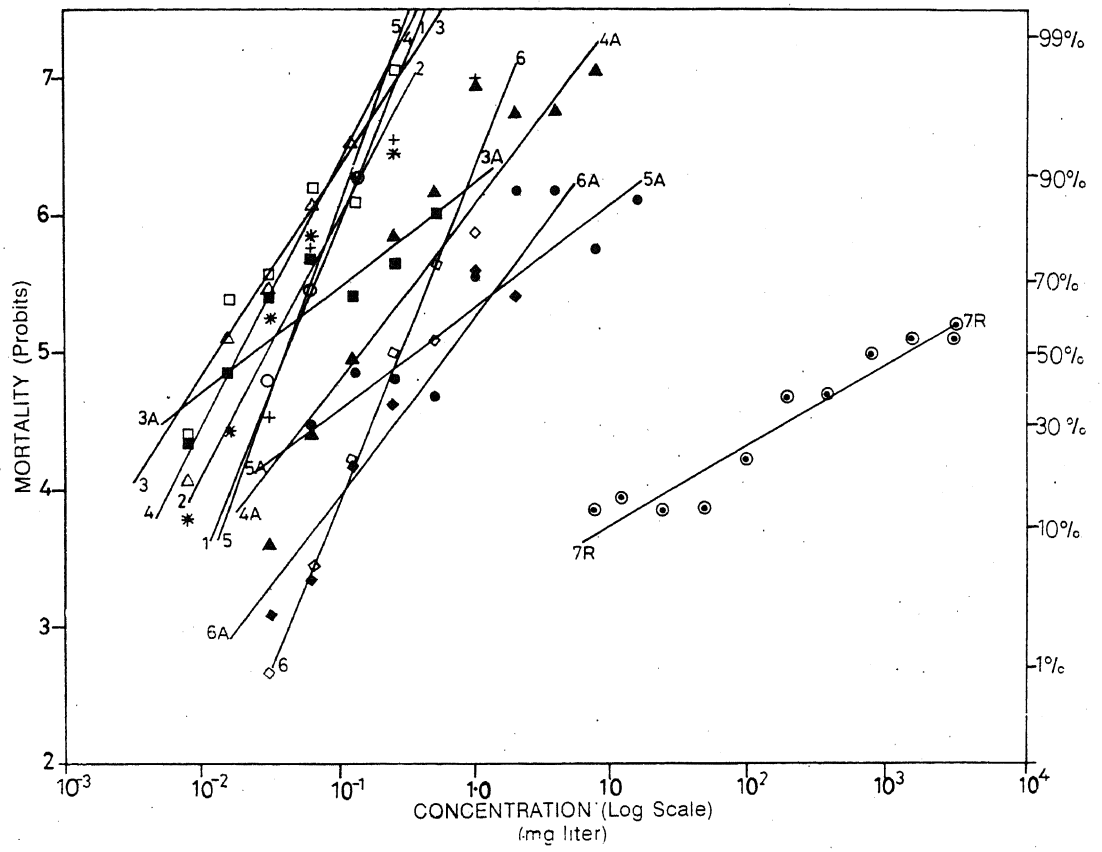


Figure 1. Dosage–mortality responses of susceptible laboratory strains from Delhi (1) and Sonepat (2) unsprayed (3–6) and sprayed (3A–6A) field populations: Farrukhabad (3,3A), Ghaziabad (4,4A), Madras (5,5A), Panaji (6,6A), and of homozygous *B. sphaericus*-resistant strain isolated from *B. sphaericus*-sprayed area of Ghaziabad (7R). Each point represents observed mortality after 24 h of treatment. At each dose 2–4 replicates (25 larvae/replicate) were used.

The LC_{50} and LC_{90} values in mg/l for Delhi and Sonepat strains were 0.04 and 0.032, and 0.134 and 0.137 respectively. These values are considered as those of susceptible populations, since these lines were established from areas where *B. sphaericus* was never sprayed. From the dose–mortality response, a dose of 2 mg/l, which was approximately double the value of $LC_{99.9}$ of susceptible strains, was calculated and subsequently used as diagnostic dose in the present studies.

The LC_{50} and LC_{90} values of the unsprayed field populations were similar to the two susceptible strains except those of Panaji, which were higher (Table 1). The higher LC_{50} and LC_{90} values in Panaji unsprayed areas could be due to the possible vigour tolerance as indicated by nonsignificant chi-square values tested for heterogeneity of response and also by the slope values which were similar to other unsprayed areas (Table 1, Figure 1). However, the LC_{50} and LC_{90} values of populations from unsprayed areas were significantly different from those of the populations from the sprayed areas. This was supported by the nonoverlapping 95% fiducial limits of LC_{50} and LC_{90} values[8] as given in Table 1. Of the four *Cx. quinquefasciatus* populations collected from *B. sphaericus* sprayed areas, chi-square values of the dose–mortality responses of Madras and Ghaziabad populations were significant ($P < 0.05$), indicating the heterogeneous nature of these populations (Table 1).

Data on LC_{50}, LC_{90}, slope values and resistance ratios as given in Table 1 clearly indicate the selection of resistance to *B. sphaericus* in field population of *Cx. quinquefasciatus*. A homozygous *B. sphaericus*-resistant strain of *Cx. quinquefasciatus* was isolated from a population collected from the sprayed areas of Ghaziabad. This strain is being maintained in the laboratory without further selection pressure. The LC_{50}, LC_{90} and slope values of the selected resistant strain in the F_6 generation were 1663 mg/l, >3200 mg/l and 0.5794, respectively (Table 1). The LC_{50} value of this strain was ~52000-fold higher than the value of the susceptible Sonepat strain. The resistant strain continued to remain resistant in subsequent generations, indicating the selection of a pure homozygous-resistant *Cx. quinquefasciatus* line. This homozygous-resistant line was subsequently used in genetic crosses.

To establish the genetic basis of resistance, the homozygous-resistant strain was reciprocally crossed with susceptible Sonepat strain. The LC_{50} (0.031 and 0.032 mg/l), LC_{90} (0.12 and 0.14 mg/l) and slope (2.26

Table 1. Regression analysis of dose–mortality responses of laboratory and field-collected strains of *Culex quinquefasciatus* from unsprayed and *Bacillus sphaericus* sprayed areas to *Bacillus sphaericus*

Strain	LC_{50}(mg/l) (–)	RR	LC_{90}(mg/l) (–)	RR	Slope	X^2 (df)
Delhi: Lab-colonized	0.04 (0.03 – 0.05)	–	0.134 (0.1 – 0.177)	–	2.46	5.76 (4) n.s.
Sonepat: Lab-colonized	0.032 (0.028 – 0.037)	–	0.137 (0.098 – 0.19)	–	2.03	7.45 (6) n.s.
Farrukhabad: Unsprayed	0.015 (0.011 – 0.021)		0.095 (0.069 – 0.132)		1.62	9.04 (5) n.s.
B.s. sprayed	0.027 (0.016 – 0.046)	1.8×	1.028 (0.36 – 2.88)	10.8×	0.81	10.48 (5)**
Ghaziabad: Unsprayed	0.02 (0.017 – 0.026)		0.09 (0.06 – 0.12)		2.08	6.39 (6) n.s
B.s. sprayed	0.14 (0.13 – 0.144)	7.0×	1.4 (0.97 – 2.03)	15.5×	1.28	29.43 (8)*
Madras: Unsprayed	0.04 (0.03 – 0.048)		0.116 (0.09 – 0.151)		2.68	2.03 (4) n.s.
B.s. sprayed	0.33 (0.22 – 0.51)	8.25×	18.0 (8.4 – 38.6)	155.2×	0.74	19.82 (7)*
Panaji: Unsprayed	0.26 (0.24 – 0.3)		0.86 (0.7 – 1.06)		2.5	4.86 (5) n.s.
B.s. sprayed	0.62 (0.5 – 0.76)	2.38×	5.97 (3.94 – 9.04)	6.94×	1.3	12.31 (6)**
Ghaziabad: Homozygous resistant	1663 (820 – 3372)	52,000×	269277	–	0.5794	6.14(8) n.s.

B.s. : *Bacillus sphaericus.*

RR : Resistance ratio = $\frac{LC_{50} \text{ or } LC_{90} \text{ of } Cx.quinquefasciatus \text{ from } B.s.\text{–sprayed area}}{LC_{50} \text{ or } LC_{90} \text{ of } Cx.quinquefasciatus \text{ from unsprayed area}}$

d.f. : Degrees of freedom (K–2).
n.s. Not significant.
* : $P < 0.05$.
** : $0.1 > P > 0.05$.
(–)95% fiducial limits.

and 2) of F_1 progeny from the reciprocal crosses were almost similar to those of the susceptible strain and all the progeny died within 24 h at the diagnostic dose (2 mg/l), indicating the recessive nature of *B. sphaericus* resistance.

In backcrosses with susceptible strain, mortalities ranged between 55.68 and 74.4%, which are well below the expected 100% at the diagnostic dose; however, mortalities reached 90–97% when scored after 48 h (after 24 h of treatment, larvae were kept in normal water). Data from crosses of F_1 progeny with resistant strain and F_1 inbreeding are given in Table 2. In backcrosses with resistant strain (crosses 1, 2, 4 and 5), mortalities were significantly below the expected 50%, and in both the F_1 inbreeding (crosses 3 and 6), dead and alive larvae deviated significantly from the expected 3:1 ratio. In all the crosses delayed effect of the larvicide was noticed. It is not clear why there was delayed mortality in all crosses, when no delayed effect of *B. sphaericus* was observed in the homozygous-resistant line and 100% mortality was observed among F_1 and susceptible progeny within 24 h of treatment. Data from the above crosses do not fit with the hypothesis that resistance to *B. sphaericus* in *Cx. quinquefasciatus* is monofactorial.

Table 2. Results from backcrosses and F_1 inbred crosses of *Cx. quinquefasciatus* strains resistant and susceptible to *B. sphaericus*

♀	♂	No. of larvae treated (no. rafts)	No. of larvae died (expected no.†)	X^2 value (d.f. = 1)
R/S	R/R	975 (7)	426 (487.5)	14.50*
R/R	R/S	1401 (10)	436 (700.5)	199.74*
R/S	R/S	994 (10)	621 (745 5)	83.29*
S/R	R/R	1431 (10)	448 (715.5)	200*
R/R	S/R	1306 (10)	492 (653)	79.6*
S/R	S/R	1663 (9)	1115 (1247.25)	56.02*

R: Resistant strain isolated from Ghaziabad.
S: Susceptible strain from Sonepat.
†Based on the hypothesis that resistance is monofactorial, expected number of larvae that would die at the diagnostic dose of 2 mg/l for 24 h.
*Significant at $P < 0.001$; ♀ = female ♂ = male.

In both F_1 inbreeding and backcrosses, among the larvae that survived, males and females were in 1:1 ratio, which indicated the linkage of the resistant gene to autosomes. In this context it may be mentioned that in *Cx. quinquefasciatus*, sex karyotype of males and females is the same, and the sex is determined by a pair of alleles, *m* and *M*. Females are homozygous for *m* (*m/m*) and males heterozygous (*m/M*). Though the F_1 progeny from the reciprocal crosses exhibited similar responses, to rule out sex linkage, surviving larvae from backcrosses and F_1 inbreeding were examined for sex by dissecting the reproductive organs. If the gene is sex-linked in F_1 inbreeding and in backcrosses, where F_1 males are heterozygous for the resistant gene, depending upon whether the resistant gene is introduced into the crosses along with *M* or *m* gene, the sex of the surviving progeny would respectively be either male or female. However, the recombinants would include the other sex. As in all the crosses males and females are in 1:1 ratio, linkage of resistant gene to sex chromosomes is ruled out.

The study has thus established that the resistance to *B. sphaericus* in *Cx. quinquefasciatus* is recessive, linked to autosome/s and is controlled by more than one gene. No maternal effect was observed in the expression of resistance. Further, the homozygous resistant strain was found fully susceptible to *B. thuringiensis israelensis*, the other commonly used biolarvicide, the LC_{50} and LC_{90} values being similar to those of the susceptible strain.

It is well established that a rare dominant gene is selected more rapidly than a recessive gene[9]. Since the gene for resistance in *B. sphaericus* is recessive, theoretically it should have taken a long time to develop resistance in the field populations. However, resistance appeared shortly after a year of spraying (20–25 rounds) in all sprayed areas. The probable reason for this could be that during larviciding, large population of immatures are continuously exposed, increasing the selection pressure[10].

Contrary to the presumption that insects are unlikely to become resistant to microbial pesticides[11], resistance in lepidopterans to *B. thuringiensis*[12] and now resistance in the mosquito *Cx. quinquefasciatus* to *B. sphaericus* has developed.

1. World Health Organization, Informal consultation on the development of *Bacillus sphaericus* as microbial larvicide, TDR/BCV/Sphaericus/85.3, 1985, 1–24.
2. Mittal, P. K., Adak, T., Batra, C. P. and Sharma, V. P., *Indian J. Malariol.*, 1993, **30**, 81–89.
3. Yadav, R. S., Sharma, V. P. and Upadhyay, A. K., *J. Am. Mosq. Control Assoc.*, 1995 (in press).
4. Anonymous, Progress report of biolarvicides in vector control, Malaria Research Centre, Delhi, 1993.
5. Singh, K. R. P., Patterson, R. S., LaBrecque, G. C. and Razdan, R. K., *J. Commun. Dis.*, 1975, 7, 31–53.
6. Mittal, P. K., Adak, T. and Sharma, V. P., *Indian J. Malariol.*, 1993, **30**, 37–41.
7. Finney, D. J., *Probit Analysis*, Cambridge University Press, Cambridge, 1971, 3rd edn, pp. 333.
8. Tabashnik, B. E., Schwarttz, J. M., Finson, N. and Johnson, M. W., *J. Econ. Entomol.*, 1992, **85**, 1046–1055.
9. Folconer, D. S., *Introduction to Quantitative Genetics*, Longman, London, 1977.
10. Forgash, A. J., *Pestic. Biochem. Physiol.*, 1984, **22**, 178–186.
11. Becker, N. and Margalit, J. in *Bacillus thuringiensis, An Environmental Biopesticide: Theory and Practice* (eds Entwistle, P. F., Cory, J. S., Bailey, M. J. and Higgs, S.) Wiley, New York, 1993, pp. 147–170.
12. Tabashnik, B. E., *Annu. Rev. Entomol.*, 1994, **39**, 47–79.

ACKNOWLEDGEMENTS. We are grateful to Drs I. Kar (Madras), A. Kumar (Panaji) and S. N. Sharma (Farrukhabad) for sending field-collected mosquitoes. We express our sincere thanks to the following staff members for their excellent technical assistance: Pritam Singh, K. B. Masiwal, H. D. Joshi, A. K. Mukherjee, Uday Prakash, Krishan Gopal, Ram Kanwar and Bhagwan Das.

Received 7 April 1995; revised accepted 4 September 1995.

Lead impairs hepatic type I-5′-monodeiodinase activity and thyroid function in cockerels

S. S. Chaurasia, P. Gupta, A. Kar and P. K. Maiti

School of Life Sciences, Devi Ahilya University, Vigyan Bhawan, Khandwa Road, Indore 452 001, India

Chronic exposure of cockerels to 1.5 mg/bird/day of lead nitrate for a period of 30 days impaired thyroid function. Lead decreased serum triiodothyronine (T_3) concentrations, inhibited hepatic type I-5′-monodeiodinase enzyme activity and marginally increased serum thyroxine (T_4) concentrations. Our findings demonstrate the lead-induced inhibition of hepatic type I-5′-monodeiodinase enzyme activity in avian system, leading to decreased production of T_3, the most potent metabolic hormone.

TYPE I-5′-monodeiodinase (5′-D) is a microsomal membrane-bound enzyme[1]. The most potent thyroid hormone, triiodothyronine (T_3), is predominantly produced in extrathyroidal tissues by 5′-monodeiodination of phenolic ring of thyroxine (T_4), the hormone that is known to be the prohormone of the former[2]. Although 5′-D enzyme system is present in almost all tissues, liver and kidney are known to have the highest T_4 to T_3 conversion activity[3]. Regulation of the 5′-monodeiodination in pe-

ripheral tissues is an important control point with respect to thyroid hormone action[4].

Lead is being increasingly recognized as an important environmental contaminant[5,6]. Many animal studies have been carried out on the toxic effects of this metal, particularly on hemopoietic, nervous, renal, cardiovascular and reproductive systems[7–10]. However, its effect on thyroid function has received very little attention of the scientists. That too, they are mostly based on simple thyroidal histology and ^{131}I uptake studies in mammal and fish[11,12]. Until today, not a single report is available on the lead-induced alteration of 5′-monodeiodination in any animal model. In the present investigation an experiment was conducted to reveal lead-induced alteration in hepatic 5′-D activity, if any, in cockerel. In addition, changes in serum T_3 and T_4 concentrations were also studied to reveal the thyroid dysfunction.

One-day-old broiler chicks were purchased from a local supplier and were acclimatized to standard laboratory conditions (14L:10D, 27 ± 1°C) for seven days. They were provided with commercial grower feed and water *ad libitum* throughout the experiment.

The birds were divided into two groups of eight birds each. Group I birds, receiving 0.1 ml of distilled water served as control. Group II birds were administered lead nitrate (1.5 mg/bird/day) subcutaneously and the treatment was continued for 30 days. The dose and the number of days were chosen from the earlier studies[13–15]. On the last day, blood was collected and the serum stored at –20°C for the estimation of T_3 and T_4 concentrations. They were estimated by the radioimmunoassay (RIA) method as described earlier, with little modification[16]. RIA kits were purchased from Bhabha Atomic Research Centre, Bombay. The lowest detection limit for T_3 assay was 0.05 ng/ml and for T_4 0.01 ng/ml. Interassay coefficient of variation was less than 5% for both the hormones.

5′-D was estimated by the method of Decuypere *et al.*[17] with little modification. The birds were sacrificed and the liver of each animal was removed and homogenized in 4 volumes (wt/vol) of phosphate buffer (0.15 M, pH 6.5, with 0.25 M sucrose and 5 mM EDTA). After centrifugation at 2000 *g* for 30 min at 4°C, the supernatant was incubated with T_4 (400 µM) and dithiothreitol (DTT, 4 mM) for 60 min. The reaction was stopped by the addition of 95% ethanol. The amount of T_3 generated was measured by RIA. Generation of T_3 was expressed as ng of T_3 generated per hour of incubation per mg protein. Liver protein was estimated by the method of Lowry *et al.*[18].

The data were expressed as mean ± SE. Significance level between two groups was calculated using Student's *t* test.

The results (Figure 1) clearly indicate lead-induced inhibition of hepatic 5′-D activity. Lead also decreased serum T_3 and marginally increased serum T_4 concentrations.

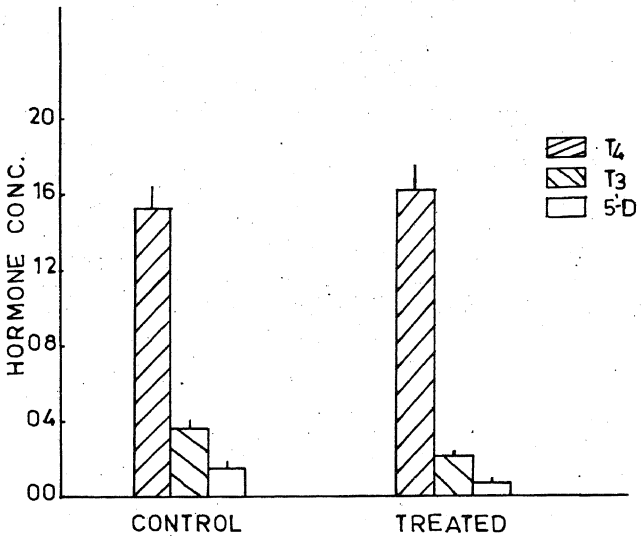


Figure. 1. Effects of lead nitrate (1.5 mg/bird/day) for 30 days on serum T_3, T_4 concentration (ng/ml) and on hepatic type I-5′-D enzyme activity (ng T_3 generated/h/mg protein) in cockerel. Vertical lines indicate the standard error of the means.

Earlier report in rat also shows a slight increase in serum T_4 concentration but a significant decrease in serum T_3 following the treatment of another heavy metal, i.e. cadmium[19]. Lead is known to induce inhibition of thyroidal iodine uptake in rats[20] and in adult patients[21]. Although at present, few reports are available on the effects of lead on thyroid function in mammals[22,23], not a single report is available on lead-induced alteration in 5′-D activity in any of the animal models. Particularly in birds, practically nothing is known on the lead-induced alterations in thyroid metabolism.

In the present study, a marginal increase in serum T_4 and a significant decrease in serum T_3, the most potent metabolic hormone, was observed following lead treatment in cockerel. The marginal increase in the serum T_4 concentration may be due to the inhibited 5′-monodeiodination (catabolism) of T_4 to T_3. This increase was not significant probably because T_4 was utilized in the other metabolic pathways[24–26]. The lead-induced decrease in serum T_3 concentration was further supported by the inhibition of 5′-D enzyme activity that is responsible for the conversion of T_4 to T_3. Earlier reports on cadmium- and mercury-induced inhibition of 5′-D in rat also support our findings[19,27].

Type I-5′-D is a sulphydryl (–SH) group-bearing enzyme and contains selenocysteine at its active site[28]. Lead has been found to inhibit a number of enzymes by either binding to its active site or to the free –SH groups of the enzyme[29,30]. Therefore, in the present study two possibilities exist for the metal-induced inhibition of 5′-D enzyme activity. Either lead replaces selenium in the enzyme or it combines with the free –SH groups, which are essential for the enzyme function, thus making it inactive.

In the present experiment, whatever may be the mechanism of action of lead on hepatic type I-5′-D enzyme, our findings for the first time reveal clearly lead-induced inhibition of the aforesaid enzyme in an avian model.

1. Sakame, S. and Chopra, I. J., *Endocrinology*, 1990, **127**, 2709–2715.
2. Chopra, I. J., Soloman, D. H., Chopra, U., Wu. S. Y., Fisher, D. A. and Nakamura, Y., *Rec. Prog. Horm. Res.*, 1978, **34**, 521–567.
3. Kuhn, E. R., Decuypere, E., Huybrechts, L. M. and Darras, V. M., in *Endocrinology of Birds: Molecular to Behavioural* (eds Wada, M. *et al.*), Japan Sci. Soc. Press. Tokyo/Springer, Berlin, 1990, pp. 129–142.
4. Kaplan, M. M. and Utiger, R. D., *J. Clin, Invest.*, 1978, **61**, 459–471.
5. *Environmental Health Criteria. 3. Lead.* World Health Organization, Geneva, 1977.
6. Lockitch, G., *Clin. Biochem.*, 1993, **26**, 371–381.
7. Scheuhammer, A. M., *Toxicology*, 1987, **45**, 165–175.
8. Jordan, S. A. and Bhatnagar, M. K., *Bull. Environ. Contam. Toxicol.*, 1987, **44**, 623–628.
9. Piasek, M. and Kostial, K., *Bull. Environ. Contam. Toxicol.*, 1987, **39**, 448–452.
10. Silbergeld, E. K. and Alder, H. S., *Brain Res.*, 1978, **148**, 451–467.
11. Katti, S. R. and Sathyanesan, A. G., *Ecotoxicol. Environ. Safety*, 1987, **13**(I), 1–6.
12. Shrivastava, V. K., Katti, S. R. and Sathyanesan, A. G., *Bull. Environ. Contam. Toxicol.*, 1987, **38**, 981–984.
13. Gochfield, M. and Burger, J., *Environ. Res.*, 1989, **50**, 262–268.
14. Nakagawa, K., *Jpn. J. Pharmacol.*, 1989, **51**, 173–179.
15. P'an, A. Y. and Kennedy, C., *Environ. Res.*, **48**, 238–247.
16. Brown, B. L., Ekins, R. P., Ellis, S. M. and Reilk, W. S., *In Vitro Procedures with Radioisotopes in Medicines*, IAEA, Vienna, 1970, pp. 569–573.
17. Decuypere, E., Scanes, C. G. and Kuhn, E. R., *Horm. Metab. Res.*, 1983, **15**, 233–236.
18. Lowry, O. H., Rosebrough, N. J., Farr, A. L. and Randall, R. J., *J. Biol. Chem.*, 1951, **193**, 265–275.
19. Yoshida, K., Sugihira, N., Suzuki, M., Sakurada, T., Saito, S., Yoshinaga, K. and Saito, H., *Environ. Res.*, 1987, **42**, 400–405.
20. Sandstead, H. H., *Proc. Soc. Exp. Biol. Med.*, 1967, **124**, 18–20.
21. Gennart, J. P., Bernard, A. and Lawerys, R., *Arch. Occup. Environ. Health*, 1992, **64** (1), 49–57.
22. Sougens, H., Klages, K., Bertrum, H. P. and Mueller, C., *Trace Element Med.*, 1987, **4**(1), 8–12.
23. Robins, J. M., Cullen, M. R., Connors, B. B. and Kayne, R. D., *Arch. Intern. Med.*, 1983, **143**, 220–224.
24. Leonard, J. L. and Visser, T. J., in *Thyroid Hormone Metabolism* (ed. Hennemann, C. G.), Dekker, New York, 1989, pp. 189–229.
25. Silve, J. E. and Leonard, J. L., *Endocrinology*, 1985, **116**, 1627–1635.
26. Visser, T. J., Fekkes, D., Docter, R. and Hennemann, G., *Biochem. J.*, 1978, **174**, 221–229.
27. Visser, T. J., van der Does–Tobe, I., Docter, R. and Hennemann, G., *Biochem. J.*, 1976, **157**, 479–482.
28. Greenspan, F. S. and Baxter, J. D., *Basic and Clinical Endocrinology*, Appleton and Lange, USA, 1994, pp. 160–226.
29. De, A. K., *Environmental Chemistry*, Wiley Eastern, New Delhi, 1990, 2nd edn, pp. 72–75.
30. Vallee, B. L. and Ulmer, D. D., *Annu. Rev. Biochem.*, 1972, **41**, 91–128.

ACKNOWLEDGEMENT. We thank UGC, New Delhi, for financial support.

Received 29 May 1995; revised accepted 2 September 1995.

MEETINGS/SYMPOSIA/SEMINARS

Workshop on Geology and Exploration of Platinum Group, Rare Metal and Rare Earth Elements

Date: 6–7 February 1996
Place: Calcutta

Topics include: Geological settings and metallogeny of the above elements; Geochemistry and mineralogy of the ores (including analytical techniques); Exploration and exploitation, including mining and benefication.

Contact: Dr A. K. Tikku
Convener, Organizing Committee
Workshop on Geology and Exploration of Platinum Group, Rare Metal and Rare Earth Elements
C/o Geological Survey of India
27, Jawaharlal Nehru Road
Calcutta 700 016
Phone: 33-2496941/2496973/2496955

Second Asian and Oceanic Congress of Andrology

Place: Chandigarh
Date: 16–20 November 1996

Contact: Congress Secretariat
Department of Biophysics
Punjab University
P. B. No. 1204
Chandigarh 160 014
Phone: 172-541441 extn 1354
Fax: 172-541409

Speaking Minds: Interviews with Twenty Eminent Cognitive Scientists. P. Baumgartner & Sabin Payr, eds. Princeton University Press, Princeton. 1995.

The realization that living organisms are really information-processing systems and, therefore, the proper framework for the study of their behaviour is the information processing (i.e. computational) framework, has brought about a major revolution in the West in traditional academic disciplines such as psychology, linguistics, education, anthropology, and even sociology. The information-processing framework as a conceptual and theorizing framework, and computer simulation studies as laboratory techniques for evaluating partially articulated theories, have been playing essential roles in drawing several of these traditional disciplines together and giving rise to interdisciplinary programmes of far-reaching consequence. Two such programmes that are beginning to have major impacts on behavioural studies are: Cognitive Sciences Programme and Neurosciences Programme.

Gardner[1] presents a very readable and a detailed historical account of the development of cognitive science which he refers to as 'the mind's new science'. According to his account, the core disciplines constituting cognitive science are: psychology, philosophy, artificial intelligence/computer science, linguistics, neurosciences, and anthropology. Gardner's definition of cognitive science is used as a guideline by the editors of *Speaking Minds*, a book of interviews with 20 academic research scientists identified as 'eminent cognitive scientists'. All the interviewees are from the United States. Their names, disciplines and institutional affiliations are as follows.

Philosophy: Patricia & Paul Churchland (UCSD); D. Dennett (Tufts); H. Dreyfus (UCB); J. Fodor (Rutgers); J. Haughland (Pittsburgh); J. Searle (UCB).
Computer Science: A. Newell (CMU: deceased): H. Simon (CMU); S. Weizenbaum (MIT: retired); R. Wilenski (UCB); T. Winograd (Stanford); L. Zadeh (UCB).
Psychology/Biology: J. McClelland (CMU); S. Palmer (UCB); D. Rumelhart (Stanford): T.Sejnowski (UCSD).
Linguistics: G. Lakoff (UCB).
Math Logic: H. Putnam (Harvard).
Sociology: A. Cicourel (UCSD).
(Note: UC, University of California; B, Berkeley; SD, San Diego; CMU, Carnegie Mellon University)

The choice of persons to be interviewed seems to have been based as much on logistic factors such as time, travel budget, availability of interviewees at the time fixed for the visits, and so on, as on the work in cognitive science these persons are engaged in and the centrality of the role they are playing in the development of this interdisciplinary field.

'Why a book of interviews?' The editors claim 'This dynamic personal discussion seemed to us – at least in some respects – much more powerful and concrete than the more general, and often more moderate and balanced, considerations in the written publications of the same scientists'. However, there are major difficulties with an interview format of the kind the editors have chosen, especially when, as in this case, the person conducting the interview is not himself/herself a practising scientist in the field being discussed critically. The most that can be done is to opt for a somewhat gossipy approach to the issues: 'So-and-So says such–and-such, what is your view on that?' or, alternatively, 'What is your view on such-and-such a controversy?', and so on.

Issues like 'Turing test', 'the chinese–room argument', 'connectionism versus physical symbol system approach', 'whether AI has made (is making) progress or is becoming irrelevant', and so on, lend themselves readily to a format of this kind – especially where the interviewees are academic philosophers with interest in AI and cognitive science. 'Listening' to arguments (gossip?) at a personal level on such issues is illuminating (or even fun) only to the extent the persons involved in the dialogue are interesting personalities in themselves. Unfortunately, this is not always the case and the dialogue pretty soon becomes rather repetitive and tiresome (especially if the personal views being thrown around are already well known to the practioners in the field). To outsiders, quite often, the discussions are likely to be of no value unless they are already 'literate' about the issues being argued about. The editors have tried to help by appending a reasonably detailed glossary of terminology. Another helpful gesture is a reading list of his/her own publications suggested by each interviewee.

In my opinion the participants in the discussions have thrown away a good opportunity to 'educate' aspiring entrants to the field of cognitive science, as well as outsiders, about the really *deep* problems that need to be addressed if cognitive science is to make any substantive contribution to our understanding of the agentive aspects of agents and how the brain/mind plays its central integrating role to enable agents to behave the way they do. What are the deep technical (i.e. computational) and theoretical (*not* philosophical) issues that need to be addressed in understanding vision, language behaviour, perception, motor behaviour (e.g. object manipulation, navigation, controlling body movements. maintaining balance, etc)? What new issues arise when behaviour involving two or more of these modalities has to be deployed to achieve some given end? Babies, only a few months old, have been shown to be capable of imitating facial expressions they are exposed to. How do they achieve this (especially since they cannot watch their own faces)? We still do not know how to construct a robot that can watch somebody else (or another robot) going through an assembly operation and imitate it. Young children are capable of accomplishing this. What are the computational issues involved here?

There are some discussions in the book about how to build up interdisciplinary programmes and why such programmes are generally more successful in the United States than elsewhere – Europe, for example. These discussions are of some relevance in the Indian context, in that in our universities (and elsewhere too) the disciplinary walls continue to remain as thick and impenetrable as ever. It may be useful to close this review with the following remarks by Patricia Churchland on why she chose to go to UCSD:

'The cognitive science community here . . . is large, diverse and lively. At some point or other, everybody passes through, so there is lots of contact, and that means there is opportunity to learn

a lot and keep abstract of new developments. There is also considerable interest and activity in computational neuroscience, so I feel much at home. I am an adjunct at the Salk Institute, and I currently work in . . . Sejnowski's laboratory. The lab has tea every afternoon, and various people drop by – often Francis Crick, and some of the visual psychologists, such as Ramachandran – and we discuss everything from consciousness and free will to apparent motion to NMDA receptor.'

1. Gardner, H., *The Mind's New Science: A History of the Cognitive Revolution*, Basic Book, New York, 1985.

R. NARASIMHAN

C/o CMC Ltd.,
Mithra Tower, 10/3 Kasturba Road,
Bangalore 560 001, India

Development of Ideas in Physics. Nils Ryde. Almqvist & Wiksell International, P. O. Box 4627, Alsnogatan 7, S-116, 91, Stockholm, Sweden. 196 pp. Price: Not known.

The rapid strides made during the first half of this century and a little earlier, have been the subject of numerous books in physics. Indeed, for several decades now, an introductory description of this area commonly known as 'Modern Physics' is an integral part of undergraduate physics curriculum in every university. It includes topics such as the discovery of electron and atomic nucleus, optical and X-ray spectroscopy, discovery of electron spin, neutron, nuclear fission and fusion and introductory quantum mechanics. The well-known textbook by Richtmyer, Kennard (and Cooper in later editions) gives an excellent exposition of these topics. Recent textbooks on modern physics like the one by Kenneth Krane also include more contemporary topics such as nucleosynthesis in stars, quark model of elementary particles and so on.

This book is a collection of a dozen partly didactic and partly historical essays on topics in modern physics. The author has been a participant in the events that unfolded in the 'golden age of physics', and his familiarity with them comes through clearly in his writing. His prose is direct and unostentatious. Most of the essays are well written, especially the ones on 'optical spectroscopy' and 'origin of the elements'. Some, however, are sketchy and need additional material for completeness (for example, the essay on matter and antimatter). The narration could have been made much more interesting with some illustrations and biographical sidelights on the main characters of the story. The quality of production is good, though a few errors remain. (For example, eq. (12) on p. 37 and eq. (21) on p. 44 contain printing errors.)

The main limitation of the book is that the essays are too 'dense' in the subject matter to be comprehensible to a beginning student. For example, in the space of some 15 pages, the author describes the rise of quantum mechanics from Heisenberg's 'arrays', deBroglie's matter waves to the successful QED explanation of the Lamb Shift in hydrogen. Such a compact description is obviously suitable only for those who are already broadly familiar with the theme. Despite this, the book does fill a certain need, since the more authoritative and complete historical accounts of modern physics (such as *The Conceptual Development of Quantum Mechanics* by Max Jammer and *Inward Bound* by Abraham Pais) are too expansive to be read with ease. The book provides short and useful historical summaries of several interesting topics, which teachers and physicists can profitably use for enrichment and recapitulation of what they already know.

ARVIND KUMAR

Homi Bhabha Centre for
Science Education (TIFR),
V. N. Purav Marg, Mankhurd,
Bombay 400 088, India

A Modern Introduction to Ancient Indian Mathematics. T. S. Bhanu Murthy. Wiley Eastern Ltd., New Delhi, 1992, 214 pp. Price: Not mentioned.
Indian Mathematics and Astronomy. S. Balachandra Rao, Jnana Deep Publications, 2388, 13th Main, A-Block, Rajajinagar, II Stage, Bangalore 560 012, India. 1994. 234 pp. Price: Rs 75, $17.00.

A man is known by the company he keeps. I suppose a similar statement holds for books as well. One of the books under review (*A Modern Introduction to Ancient Indian Mathematics*) suffers because of its unqualified association with Swami Bharati Krishna Tirtha's *Vedic Mathematics*. The author mentions in the preface that the initial purpose of the book was to present proofs of certain propositions stated in *Vedic Mathematics*, thus subtly authenticating it. (It is more an error of omission than of commission as he fails to warn the readers about the doubts expressed by scholars in respect of many of Swamiji's claims.) First, a few words about *Vedic Mathematics*. It is more or less established that its contents are not of vedic origin (the Swamiji has failed to produce a reliable reference) and its mathematical contents do not, by a long chalk, justify the tall claims made in the book. Take the example of the algorithm to write down the decimal expansion of reciprocals of certain integers. As has been pointed out by S. G. Dani in his excellent two-part article ('Myths and Reality – On Vedic Mathematics', *Frontline*, 22 October 1993 and 5 November 1993), there is no evidence in the available vedic works or even in the works of other Indian mathematicians for over 2000 years later, of any knowledge of writing a fraction in decimal form. Unfortunately, this is one of the propositions Bhanu Murthy chooses to prove in his book.

Apart from these proofs (of Swamiji's propositions), Bhanu Murthy's book contains a well-written account of the *Pell's* equation (which he prefers to refer to as Brahmagupta–Bhaskara equation). There is also a chapter entitled 'Selected Topics in Geometry' and the reviewer fails to understand why some of the topics which are featured are there at all; for instance, Morley's theorem [which would have had some meaning at least if he had given the elegant proof due to M. T. Naraniengar in 1909 (see *Geometry Revisited*, Coxeter and Greitzer)].

On the other hand S. Balachandra Rao's book, *Indian Mathematics and Astronomy*, provides a good introduction to the development of mathematics and astronomy in India since vedic times. Each mathematician's contribution is preceded by a discussion on his place and times according to available records. In fact, one of the praiseworthy features is that statements made are justified with references and different viewpoints are sincerely reproduced. There is an excellent bibliography at the end.

One of the main points that Balachandra Rao (as also Bhanu Murthy) wants to emphasize is that it is not true, as is generally believed, that Indian mathematics entered a dark age after Bhaskara II (*ca* 1120 AD). There were substantial contributions even after him, specially by the Kerala school. In fact, to quote Dani '... The sustained development of mathematics in India in the post-Greek period was indirectly instrumental in the revival in Europe after its "dark ages" '.

Though the achievements of mathematicians in ancient India may look impressive, there were serious deficiencies also. They have noticed that π cannot be calculated exactly but nothing seems to have been done to investigate this aspect further in contrast to the amount of time and effort that have gone into *squaring the circle* in the Greek civilization. Indian mathematicians have handled surds with ease but the notion of irrational numbers escaped them. The deductive reasoning which characterized the Greek mathematics and which really gave the impetus for the phenomenal development of mathematics outside India is conspicuous by its absence in Indian mathematics. Another significant factor which impeded the progress was that mathematics mostly served as the hand-maid of astronomy whose needs it soon outgrew. One also has to consider the fact that physics had not developed much, since it was in trying to answer the needs of physics that calculus was discovered and several other areas of mathematics developed. As Balachandra Rao rightly observes, the perfection to which Indian logic (tarka) was developed seems to have had no influence on the thinking of ancient Indian mathematicians. To quote him further: 'The over-emphasis on utilitarianism of mathematics seems to have had unfortunate results in the development of the subject in this country... The nonchalance with which the splendid achievements of the Greek geometry were ignored while the pseudoscience of Greek and Babylonian astrology was received with open arms is indeed the worst of these.'

For further serious study of history of mathematics in India it is very much necessary to bring out authoritative translations of the original works. For most of them the translations available are more than 60 years old and so do not have the benefit of research done on the subject in recent times, which is quite substantial. Also, some positive attempts should be made to educate the public about the facts concerning *Vedic Mathematics*. It has acquired a large following and there are even correspondence courses on the subject! There is a danger, a very real one at that, of this having undesirable consequences on mathematics education in our country. Already the Uttar Pradesh government has introduced 'vedic mathematics' in school textbooks based on the Swamiji's book. While it is important to create an awareness of our past achievements (based on authentic sources) it is also necessary to confront and expose baseless constructs such as these.

C. S. YOGANANDA

Department of Mathematics,
Indian Institute of Science,
Bangalore 560 012, India

Leks. Jacob Hoglund and Rauno V. Alatalo. Princeton University Press, 41 Williams Street, Princeton, New Jersey 08540, USA. 1995. 248 pp. Price: Not known.

Blazes of swarming fireflies in the woods, synchronized arm-waving of fiddler crabs on estuarine mudflats, incessant nocturnal choruses of crickets, katydids and frogs/toads in our neighbourhoods and dancing peacocks are familiar yearly events in India. These male aggregations are meant to attract females for the purpose of mating. However, not all these events are treated in the strict sense as 'lekking'.

Lekking has been defined as 'males aggregating for the purpose of mating'. In lekking animals the male's contribution to parental care is none, females have the option of mating with males of their choice and female gathering on the leks is not a response to some resource such as food or ideal breeding site that the male actively defends. Thus by the above definition, many mating aggregations of animals such as that of fiddler crabs, some fishes and frogs cannot be treated as typical lekking, as males of these animals are known to defend egg deposition sites simultaneously as they display.

Almost 200 years ago, naturalists observed and reported mating aggregations on natural 'arenas' in vertebrate animals. Somewhere in the late 1800s, the term 'lek' was first coined to describe the arena and 'lekking' for the rather deviant mating system. Lekking is now reported in many groups of lower organisms including insects and spiders. Nevertheless, the most typical and well-studied forms of lekking are those that the birds and a few ungulate mammals demonstrate. And probably the most intensively studied of animals is the temperate sage grouse, a galliform bird that mates on leks throughout its range.

Field study of animal behaviour has not gained the same momentum in India as it has in the western world. In India there have been some attempts to study the breeding ecology and behaviour of animals in the field starting with Salim Ali's work on the weaverbirds in the nineteen thirties. Although there have been a few casual reports of lekking in peacocks and blackbucks in the past, the only real attempt to study the ecology of lekking in Indian animals *in situ* is a recent Master of Science dissertation from the Wildlife Institute of India, Dehra Dun. This rather illuminating piece of work on lekking blackbucks in north-western India is however unpublished.

The book under review is easily the most complete treatise on this special subject relating to animal mating systems in recent times. It is written in a very simple and readable style. The authors have been considerate in keeping the more complex mathematical modelling as a latter chapter without depriving the uninitiated reader the pleasure of going through the book.

The organisms listed as observed lekking in chapter 2 are surprisingly many and spread across taxonomic groups. Taking a variety of general and detailed case studies from this spectrum of animals, the authors discuss the broad causal factors of lekking. Although detailed discussions of sexual selection, female mating strategies and a number of models that attempt to explain the advantages and disadvantages of lekking, including predation risk, sharing information, etc., are invoked, what ultimately perplexes the reader is the fact that there has not been one gen-

eral governing principle that has hitherto been identified without ambiguity.

Lekking has best evolved in birds. Interestingly, of the nearly 100 species of birds (as listed in the book) in which this behaviour has been documented/studied in detail, 78% are forest birds, especially tropical, as that in families Paradiseidae (New Guinea/ Australasia), Cotingidae and Pipridae (both Neotropical). Similarly, it is noteworthy that 62% and 20% of all lekking birds belong to the New World and Australasia respectively. These patterns are suggestive enough to start speculating the role of biogeography and habitat on the origins of this particular mating system, at least in birds.

Tropical rainforests in New Guinea and America are very dense and dark within. Male birds of paradise, cotingas and manakins with their elaborate and strikingly irridescent plumage will certainly attract females better by aggregating and displaying than being solitary. Although related ideas, viz. 'visual stimulus' and 'passive attraction' to females offered by aggregating males have been addressed and dismissed as causal factors in the book, these ideas have not been put to good test in the tropics.

Whether this behaviour is 'relict' in animals has not been addressed in the book. Considering the biogeographical and habitat-related pattern of distribution of lekking behaviour in birds, it is also tempting to invoke the argument of patchy habitat availability in the tropics during the Pleistocene glaciations. Works in the early 1970s and 1980s have discussed the possible influences that reduced availability of forest habitats had on the diversity and speciation of neotropical birds. If reduced habitat availability had forced breeding birds to aggregate and display during the long periods of climatic fluctuations, such traits could not have been eliminated during the last 10,000 more or less steadily warm years. When ecology fails to explain an observed pattern in nature, scientists have resorted to history. The book under review fails to do this.

More examples of birds, especially the migratory waders such as snipes and ruff/reeve in which lekking is reported, are compelling enough to invoke historical explanations to the origin and persistence of lekking. Migration in birds is itself considered a climate change-induced relict behaviour. Future studies should probably look at the lower animals, especially the invertebrates in the light of the existing theories on lekking, not ignoring the importance of history. Any further knowledge on animal mating systems will not only be of academic interest but of great value in the conservation of biodiversity.

R. J. RANJIT DANIELS

M. S. Swaminathan Research Foundation, 3rd Cross Street, Taramani, Institutional Area, Madras 600 113, India

MEETINGS/SYMPOSIA/SEMINARS

X National Carbohydrate Conference

Date: 8–9 November 1995
Place: Vallabh Vidyanagar

Topics include: Plant and microbial polysaccharides; Isolation, characterization and their biological activity; Carbohydrate synthesis; Industrial polysaccharides: Production and utilization; Carbohydrate fermentation/biotransformation.

Contact: Prof. M. N. Patel
Convener, Carbohydrate Conference
Department of Chemistry
Sardar Patel University
Vallabh Vidyanagar 388 120
Phone: (02692) 30411
Fax: 02692-30238

National Workshop on Biodiversity Conservation in Managed Forests and Protected Areas

Date: 29 November to 1 December 1995
Place: Bhopal

Topics include: Biodiversity conservation in managed forests; Biodiversity in protected areas; Ethical and ecological considerations; Socio-economic issues; Institutional, policy and planning issues.

Contact: Prof. P. C. Kotwal
Indian Institute of Forest Management
Nehru Nagar, P. B. No. 357
Bhopal 462 003

Second National Symposium on Magnetic Resonance

Date: 8–9 February 1996
Place: Pune

The symposium will cover all topics on magnetic resonance and its applications in physics, chemistry, biology and medicine.

Contact: Dr S. Ganapathy
Solid State NMR Group
National Chemical Laboratory
Pune 411 008

Samanta Chandra Sekhar: Life and work

P. C. Naik and L. Satpathy

Mahamahopadhyaya Sri Chandra Sekhar Simha Hari Chandan Mahapatra Samanta (1835–1904), generally known as Samanta Chandra Sekhar and more popularly as Pathani Samanta in Orissa, worked in astronomy following the traditional methods and composed a treatise entitled 'Siddhanta Darpana' in Sanskrit verse. It is a valuable work that records innumerable improvements over the ancient classics, namely, 'Surya Siddhanta' and 'Siddhanta Siromani', by incorporating original observations, calculations, instrumentation and theories. The aim of this paper is to outline and bring to the notice of a wider audience the genius of Chandra Sekhar and his contribution to astronomy. In addition, we also present a brief life-sketch of the astronomer.

India has a rich heritage of astronomy having produced great astronomers like Aryabhata, Varahamihira, Brahmagupta and Bhaskara. The genius of these luminaries bloomed at various places spread all over the country, scaling a time span of nearly 1500 years beginning with 5th century AD. Mahamahopadhyaya Chandra Sekhar Simha Harichandan Mahapatra Samanta, referred to hereafter as Samanta Chandra Sekhar, seems to be the last link of this long order of great Hindu astronomers. *Surya Siddhanta*, *Aryabhatiya*, *Panca Siddhantika*, *Brahmasphuta Siddhanta* and *Siddhanta Siromani* are prominent among the numerous astronomical works of our illustrious ancestors. *Siddhanta Darpana*[1], composed by Samanta Chandra Sekhar, looks to be the latest valuable classic of high order. It is a systematic record of Samanta's lifelong research in astronomy with substantial original contribution to the field. The treatise, written in beautiful Sanskrit verse after the time-honoured Hindu tradition, has great scientific value and is also a fine piece of literary work.

Chandra Sekhar was born in the royal family of one of the princely states of Orissa, at Khandapara in 1835. He is also fondly called Pathani Samanta in the state. Chandra Sekhar, when a child, was initiated to identify stars by his father. He received primary education from a Brahmin teacher. Thereafter, he started teaching himself *Lilavati*, *Bijaganita*, *Jyotisha*, *Siddhanta*, *Vyakaran* and *Kavya* using his family library. He was more attracted by mathematics and astronomy.

At the age of fifteen, he began to check the predictions of the *Siddhantas* with his observations. Ancient Indian works do not give details of instruments or methods of measurement explicitly, excepting hints here and there. So young Chandra Sekhar devised his own instruments. He noticed that, even classics like *Siddhanta Siromani* of Bhaskara and *Surya Siddhanta* did not agree with his observations, for the stars and planets did not appear in the sky as per their predictions.

Hence, young Chandra Sekhar assigned the job to himself and set out for observations night and day, checking every figure occurring in the earlier works.



He began recording his observations and formulations at the age of 23, and in the form of treatise, two years later. The manuscript of *Siddhanta Darpana* was ready in 1869 when he was 34. But it took another 30 years for the work, originally written on palm leaves in Oriya characters, to appear in Devanagari script printed on paper in 1899. Chandra Sekhar had *Surya Siddhanta* for his prime reference and *Siddhanta Siromani* as his guide.

This remarkable scholar and scientist, being confined to a small hilly state in a remote corner of Orissa, far away from the sphere of English education, was virtually doomed to obscurity but for his chance acquaintance with Mahesh Chandra Nyayaratna, Principal of Sanskrit College, Calcutta, who probably later on introduced him to Jogesh Chandra Ray of Cuttack College (today's Ravenshaw College). Nyayaratna's efforts[2] brought him the title of Mahamahopadhyaya, conferred by the British Government in 1893.

Ray was the key person under whose supervision and involvement *Siddhanta Darpana* was published in 1899 from Calcutta with the financial support of the kings of Athmallik and Mayurbhanj. It must be noted that the scholarly introduction of 56 pages in English therein by Ray[3] formed the window through which the outside world could get a glimpse of the valuable treasures contained in this monumental work couched in hardly accessible Sanskrit verse. The international journals *Nature*[4] and *Knowledge*[5] acknowledged the greatness of this contribution to astronomy by Chandra Sekhar and eulogized him in glowing terms. This Sanskrit classic underwent two Oriya translations[6,7] in 1975 and 1976. We present in the following a brief discussion of this work citing only a few important original contributions of its author.

Siddhanta Darpana

Chandra Sekhar in the 17th *Sloka* of the first chapter of the treatise enumerates the requisite characteristics of a *Siddhanta*. The work that deals with the theoretical division of time from the smallest unit of *Truti* (0.274348×10^{-6} s)

Table 1. Summary of the contents of *Siddhanta Darpana*

Name of section	Chapter number	Chapter content	Number of *Slokas* in the chapter
Madhymadhikara	1	Description of time	55
	2	Description of *Bhagan*, etc.	25
	3	Mean planet positions	77
	4	Various corrections	57
Sphutadhikara	5	True planet positions	221
	6	Finer corrections	161
Triprasnadhikara	7	Gnomons, etc.	94
	8	Lunar eclipse	87
	9	Solar eclipse	78
	10	Parilekha description	37
	11	Transit, etc., of planets	111
	12	Alignments of planets	93
	13	Risings and setting of planet	84
	14	Phases of the moon	67
	15	Description of Mahapata	70
Goladhikara	16	A set of questions	79
	17	Description of the Earth	159
	18	Description of the Earth (contd.)	175
	19	The celestial sphere	123
	20	Description of instruments	111
	21	Some deeper questions	249
Kaladhikara	22	Description of years, etc.	76
	23	Prayer to Purusottama	55
	24	Conclusion	156
Total			**2500**

to *Pralaya* (10^{13} solar years), along with the motion of the celestial objects, their rotations, orbits, alignments, occultation and eclipses, etc., with the relevant mathematics like algebra, arithmetic, geometry and trigonometry and also concerns with the question of the origin of the universe is called a *Siddhanta*. One will notice that *Siddhanta Darpana* with 24 chapters and 2500 *Slokas*, out of which 2284 are Samanta's own composition and 216 are citations from the ealier authors, falls short of none of these qualifications. Chandra Sekhar has made original contributions to most of the aforesaid topics dealt with in his astronomical treatise.

It will not be out of place here to give a brief contentwise description of *Siddhanta Darpana*. It is broadly divided into two parts, viz. *Purvardha* (first half) and *Uttarardha* (latter half). The first half contains 15 chapters and the latter half 9. The chapters are further grouped under five sections, namely, *Madhyadhikara*, *Sphutadhikara*, *Triprasnadhikara*, *Goladhikara* and *Kaladhikara*. The first two sections deal with the mean motion and the true positions of the planets, respectively. The third section deals with motion described in terms of space, time and direction. The fourth section gives an account of the relevant mathematics like spherical trigonometry and geometry, etc. The fifth section describes different ways of reckoning time. The distribution of chapters and *Slokas* along with the title of the topic dealt with in each chapter, is given in Table 1.

The contents of *Siddhanta Darpana* look amazing as the achievement of a single mind. Chandra Sekhar has observed, verified and corrected wherever necessary all that was known to the Hindu astronomers for thousands of years. Very often he has gone beyond them to discover new phenomena and to give new formulations and comes out with predictions supposed to remain valid for at least 10,000 years to come. We give below an outline of his contributions to the four important aspects of astronomy, namely (i) observation, (ii) calculation, (iii) method of measurement and instrumentation and (iv) theory and model.

Observation

As indicated in the introduction, the discrepancy between predictions of the ancient *Siddhantas* and his own observations prompted Chandra Sekhar to embark upon serious investigation. He has made it clear that, apart from purely theoretical questions, nothing has been recorded in the *Darpana* without verification through observations.

In the third chapter of the treatise, Samanta enters into discussion on the planets. Let us look at the sidereal periods (time taken for a complete revolution with respect to 'a fixed' star) of the planets as ascertained by him. For comparison, the periods given in *Surya Siddhanta* and *Siddhanta Siromani*, the

Table 2. Sidereal periods of the planets in mean solar days. The data in the last column are taken from ref. 19 and the remaining ones from refs 3 and 20

Planet	*Surya Siddhanta*	*Siddhanta Siromani*	*Siddhanta Darpana*	European values as in 1899	Modern values
Sun	365.25875	365.25843	365.25875	365.25637	365.25636
Moon	27.32167	27.32114	27.32167	27.32166	27.3216615
Mars	686.9975	686.9979	686.9857	686.9794	686.97982
Mercury	87.9585	87.9699	87.9701	87.9692	87.969256
Jupiter	4332.3206	4332.2408	4332.6278	4332.5848	4332.589
Venus	224.6985	224.9679	224.7023	224.7007	224.70080
Saturn	10765.7730	10765.8152	10759.7605	10759.2197	10759.23

Table 3. Inclination of the orbits of planets to the ecliptic. The data in the last column are taken from ref. 19 and the remaining ones from refs 3 and 20.The quantities are in degree, minute and second

Planet	*Surya Siddhanta*	*Siddhanta Siromani*	*Siddhanta Darpana*	European values as in 1899	Modern values
Moon	4 30 0	4 30 0	5 09 0	5 08 48	5 08 33
Mars	1 30 0	1 50 0	1 51 0	1 51 2	1 50 59
Mercury	5 55 0	6 55 0	7 2 0	7 00 08	7 00 18
Jupiter	1 0 0	1 16 0	1 18 0	1 18 41	1 18 18
Venus	2 46 0	3 6 0	3 23 0	3 53 35	3 23 41
Saturn	2 0 0	2 40 0	2 29 0	2 29 40	2 29 10

Table 4. Greatest equations of sun and moon (see text). The data are taken from ref. 3. The quantities are in degree, minute and second

Celestial body	*Surya Siddhanta*	*Siddhanta Darpana*	Western astronomy as in 1899
Sun	2 10 30	1 55 33	1 55 19
Moon	5 2 46	5 1 10	6 3 41

European values as known in 1899, and the modern ones are presented side by side in Table 2. It may be noted that Chandra Sekhar has arrived at the same values for the periods of the sun and the moon as in *Surya Siddhanta*, which agree with the modern values up to second and third decimal places, respectively. But in the case of the slow-moving planets like Jupiter and Saturn, his results are surprisingly in closer agreement with the modern values than those of the other two *Siddhantas*. Table 3 gives a comparison of the mean inclination of the orbits of the planets to the ecliptic. It can be seen that, of the three classical works quoted here the values of *Siddhanta Darpana* are the closest to those of modern astronomy.

The angular position of a planet in its orbit, with respect to the force centre, is called the 'true anomaly'. The eccentric angle of the planet on the auxiliary circle is the 'eccentric anomaly'. Their difference is the equation of anomaly, the greatest value of which is a measure of eccentricity. The greatest equations given by Samanta in the case of the sun and the moon agree within seconds with the Western astronomy (Table 4). The intersection positions of the moon's orbit with the ecliptic are called *Rahu* and *Ketu* in the traditional Indian astronomy. In the Western astronomy they are called the nodes of moon. The minimum distance position of the moon from the earth is called the *perigee*. Table 5 shows a comparison of the sidereal periods of these nodes and the perigee.

Further, Chandra Sekhar gives a correction to be applied to *Mandocca* (the farthest position on the orbit from the force centre) of Mercury, Mars and Saturn and calls it '*Parocca correction*'. The maximum values of these corrections are 12°20′, 7°30′ and 5° for the above three planets, respectively.

In observation of the moon, Samanta Chandra Sekhar noticed three important irregularities. They are *Tungantara* (evection), *Pakshika* (variation) and *Digamsa* (annual equation). In terms of modern understanding[8,9], recognition of these three irregularities is as follows:

(i) *Tungantara* (evection) is the largest of the irregularities. It depends on the changing position of the apse line (the line joining the maximum and the minimum distance positions from the force centre, i.e. the earth) in successive lunar months. It further has the following two effects:

(a) The apse line advances as a whole while undergoing oscillatory motion. However, it retrogrades when the perigee (maximum distance from the earth) occurs at the first and third quarters of a lunar month.

(b) The eccentricity also oscillates, attaining a maximum of 0.066, when perigee occurs at new or full moon, and a minimum of 0.044, when it occurs at first or third quarter. This perturbation effectively puts the moon forward or backward from the expected position by about a degree. The maximum of this irregularity according to modern astronomy is 1°17′, whereas Samanta Chandra Sekhar observed it to be at most 2°40′.

(ii) *Pakshika* (variation): From new moon to full moon, i.e. during the bright fortnight of the lunar month, the attractive force of the sun on the moon retards it. And during the dark fortnight, the effect is opposite and the moon is accelerated. This causes a maximum shift in the position of the moon by 39′31″ according to modern astronomy. Samanta's value of 38′12″ for maximum of this effect is surprisingly close to the modern value.

(iii) *Digamsa* (annual equation): This is another irregularity in the position of the moon, which arises due to variation in sun–moon distance as the earth orbits round the sun. Obviously, its period is one year. The maximum value of this correction is 11′9″ as per modern astronomy, and Samanta's value of 12′ is fairly close to it.

Chandra Sekhar has clearly mentioned in *Siddhanta Darpana* that *Tungantara* was known to his predecessors. However, J. C. Ray in his introduction to *Siddhanta Darpana* says: 'In computing the moon's place, Chandra Sekhar has discovered some original corrections – original in the sense of their having been unknown to the ancient astronomers of our country'[3]. Even Burgess[10] in his English translation of *Surya Siddhanta* says that the Hindus did not take any notice of evection. However, the latest view on this issue is that of P. C. Sengupta[11], who concludes that evection and variation were introduced by Munjala and Bhaskara. Nevertheless, it is definite that Samanta Chandra Sekhar was the only Indian astronomer of traditional school to have discovered *Digamsa* and took into account all the three irregularities and measured them with utmost accuracy. In passing, it will not be out of place to mention that Tycho Brahe was the first

Table 5. Sidereal periods of the nodes (*Rahu* and *Ketu*) and the perigee in mean solar days (see text). The data in the last column are taken from ref. 19 and the remaining ones from refs 3 and 20

Entity	*Surya Siddhanta*	*Siddhanta Siromani*	*Siddhanta Darpana*	European values as in 1899	Modern values
Moon's nodes	6794.3948	6793.2535	6792.644	6789.270	6793.470
Moon's perigee	3232.094	3232.734	3232.657	3232.575	3232.600

Table 6. Rates of mean motion of the planets per day. The quantities in the upper line in each row in the third and fourth column are in sexagesimal system and the ones in the corresponding lower lines are in decimal system. The data under modern value are taken from ref. 19 and those due to other *Siddhantas* from ref. 20. The quantities are in degree, minute and second

Planet	Modern value (Lahiri)	Brahmagupta Sripati Bhaskar	Samanta Chandra Sekhar
Sun	0 59 08.2	0 59/8/10/21	59/8/10/24/12/30/4/0/4
		0 59 8.1725	0 59 8.1733912182
Moon	13 10 35.0	13/10/34/53/0	13/10/34/52/03/49/08/02/16/10/11
		13 10 34.8833	13 10 34 8677275173
Mars	0.31 26.5	0/31/26/28/7	0.31 26.5018877228
		0 31 26.4686	
Mercury	4 05 32.4	4/05/32/18/28	
		4 5 32.3077	
Jupiter	0 04 59.1	0/04/59/9/9	0/04/59/05/37/0/36/41/17/01/51
		0 04 59.1525	0 04 59.0936139473
Venus	1.36 07.7	1/36/7/44/35	
		1 36 7.74305	
Saturn	0 02 01.9	0/20/0/22/51	
		0 2 0.380833	0/02/0/26/55/02/53/21/02/04/54
Moon's nodes	–0 03 10.6	0/03/10/48/20	0 2 0.448624511
		0 30 10.8055	0 03 10.7946631769

Western astronomer to have detected and measured all these three irregularities.

Another improvement made by Chandra Sekhar was in the case of parallaxes of sun and moon. Parallax is defined as the angle subtended by the radius of the earth at a heavenly body. The earlier *Siddhantas* took it to be 3′56″at sun and 52′42″at moon. But since these are the quantities that enter into calculation of eclipses, Chandra Sekhar changed them to make the predictions agree with the observations. The ancient Indian astronomers place the sun at a distance not more than 14 times that of the moon from the earth. Chandra Sekhar hiked this ratio by more than ten times, to 154, which is much closer to the modern value of 390.

Calculation

Astronomy is a science in which observation goes hand in hand with calculation. Besides a thorough knowledge of the motion of the celestial objects, it requires sufficient computational skill to tackle the vast multitude of big numbers We will see in this section that Chandra Sekhar was extraordinary in this respect. Very often he has gone beyond his predecessors in precision of computation and in giving new methods of calculation.

To get positions of planets one has to do extremely careful calculations. In the traditional method one first calculates a mean position and then applies corrections to arrive at the true position. The calculation of mean position requires a mean rate of motion. Works in traditional astronomy clearly give these rates as accurately as possible. The earlier works are found to give these rates (mean angular displacements per day) at best to five places in the sexagesimal system. But Chandra Sekhar gives these rates to ten places in the same system. For comparison, we have presented in Table 6 the modern values of these rates in the decimal system and also given the traditional data converted into decimal system.

Chandra Sekhar recognized that, in addition to the usual correction to the mean position, further finer corrections are needed from time to time to make accurate predictions. For each planet, he has provided such a correction to be applied for hundred million (one *Arbuda*) years. There are as many as 55 tables, each on an average having 50 numbers, sometimes given up to five places in sexagesimal system. These tables appear to be a challenge for one's unaided computational skill.

Regarding the measurement of longitudes of the celestial bodies, there is a fundamental difference between the Western system and that of the Hindus. In the former system, it is measured with respect to the vernal equinox, i.e. one of the point of intersection of the celestial equator with the ecliptic. The Hindu system measures it with respect to a fixed point (a particular star) on the ecliptic. It is the starting point of zodiac and its longitude is the *Ayanamsa*. But it may be pointed out that the axis of the earth undergoes a precessional motion with a period of about 25,800 years. This causes a shift in the position of the equinox and also in that of the reference star, which according to the traditional Indian astronomy is called *Ayanacalana*. So, this effect must be taken into account in fixing the *Ayanamsa* of the reference point. A correct determination of the *Ayanamsa* involves a number of factors like the rate of pre-

cession, the exact length of a year, the year of the zero *Ayanamsa* and the star of reference. A critical discussion of the matter is found in the introduction to *Siddhanta Darpana*[3]. *Surya Siddhanta, Soma Siddhanta, Sakalya Siddhanta*, and *Laghu Vasistha*, etc., have given the rate of precession as lying in the range 54″ to 46″.25 per year. However, *Bhasvati* and *Graha Laghava* have found the value to be nearly 60″. But Chandra Sekhar's value is 57.615″ per year.

The modern value of the precessional rate is, however, 50.3″ per year. Thus, it appears that the Indian rates quoted above are in disagreement with the modern value. Therefore, it is desirable to understand how such varying rates of precision which sensitively affect the fixation of *Ayanamsa* could give meaningful prediction in the ancient Indian astronomy. Ray gives a solution to this puzzle. He points out that the Indian astronomers take the sidereal year for calculation of the rate of precession, in contrast to the tropical year, taken in the West. Since the sidereal year is longer by a few minutes, the sun advances by an amount of 8.4″ during this interval of time difference. This has an effect on the calculation of precession of equinox and the amount has to be deducted from the rates given by various authors. When this correction is taken into account, Samanta's value turn out to be 49″.179 per year, which is much closer to the modern value.

Chandra Sekhar has devised a diagramatic or graphical method (*Parilekha*) for calculations of eclipses and many other phenomena. In this method, when the nature and periodicity of the motion of the heavenly bodies are known, one can simulate the phenomena by constructing geometrical diagrams on the plain surface and use them for computation of dynamical quantities. This method might be akin to the numerical evaluation of integrals. Especially, he tackles the calculations of the three well-known irregularities of the moon with simple geometrical arguments. It is needless to say that his calculations of eclipses are accurate and agree with observations. It may be noted here that this has been the guide in the preparation of the almanac used in Orissa since about 1870.

Before concluding this section, we would like to point out at least two facts of mathematical interest from *Siddhanta Darpana*. Samanta does not accept the value of π as 22/7; instead, he takes it to be 3927/1250 and also 600/191. He seems to have provided a new method for computing the cube root of numbers which is different from that of Bhaskara.

Measurement and instrumentation

Chandra Sekhar was unaware of the scientific and technological developments that had already taken place in the West. He had not seen or even heard of a telescope until his very old age. So he employed the traditional methods for astronomical measurements. He had devised some new instruments for the purpose. It is a remarkable point that Jai Singh in the 18th century built a number of observatories in the northern part of the country. His instruments for measurements in these observatories are rather gigantic. In sharp contrast, Chandra Sekhar, who worked about 150 years later, fabricated quite simple and small devices and made accurate observations.

Before attempting to give an account of his astronomical *Yantras*, we would like to mention an interesting fact about Chandra Sekhar's ability for measurements. The children at primary schools in Orissa today read it as a myth that Samanta measured the heights of distant mountains with a pair of tiny sticks. The simplest of his stick instruments is the *Sanku* (pole). Of course, it was used also by the Greeks under the name '*gnomon*'. It is not difficult to find out the latitude of a place by observing the shadow of the *Sanku*, cast by the sun. But more interesting is, Chandra Sekhar's calculation of local time from the shadow.

Chandra Sekhar's most talked of instrument is the *Mana Yantra*. Details of construction of this instrument have been described in *Siddhanta Darpana*. Chandra Sekhar used this *Yantra* as a multipurpose device. A simpler version of this *Yantra* is a two-stick arrangement in the form of 'T', which he could fabricate quickly, and measure the height and distance of mountains, etc., whenever occasion demanded.

Samanta has described in detail the construction techniques for varieties of *Gola-yantra* that he used to measure the positions of planets. There is a mention of as many as ten different devices for measurement of time in *Siddhanta Darpana*. Models of some of these, including a *Gola-yantra* have recently been exhibited at the Regional Science Centre, Bhubaneswar. Samanta Chandra Sekhar has devoted a full chapter of *Siddhanta Darpana* to describing the construction and use of instruments. Such instruments, by today's standard, are crude, but using them he could arrive at results in close agreement with modern values, primarily due to his skill, attained through long years of dedication and practice.

Theory and model

The most prominent of Chandra Sekhar's theoretical contributions is his model of the planetary system, which is different from those of his predecessors in India. Samanta supported a stationary-earth hypothesis, with sun, moon and stars revolving around the earth. But it has the novel feature that planets like Mercury, Venus, Mars, Jupiter and Saturn go around the sun, and the sun moves around the earth together with these companions. This is evident from the following *Slokas* of *Siddhanta Darpana*:

तन्नमध्यम-मार्त्तण्डः परितोमण्डलंभुवः
भ्रमन् ताराखेचराणां कक्षामध्यस्थ उच्यते।
तं भ्रमन्तो महीजाद्यास्तत्सङ्गेन भुवं पुनः
परिक्रामन्ति यत्तस्मात् स प्रोक्तः सर्वकर्षकः।

S. D. Ch. V Slokas 6 and 7

Thus, he assigned heliocentric motion to the planets. Incidentally, a similar model was proposed by Tycho Brahe in the West in the late sixteenth century. A recent work by Ramsubramanian *et al.*[12] brings to light that a similar model of the planetary system was given by Nilakantha Somasutvan of the Kerala school by 1500 AD.

Conclusion

We have attempted to give a brief sketch of the extraordinary life and work of Samanta Chandra Sekhar here. Unfortunately, due to various historical and geographical reasons, his true position in ancient Indian astronomy has not been ascertained so far. Ray has drawn a nice parallelism between the life and works of Samanta Chandra Sekhar and of Tycho Brahe. The journal *Nature* com-

ments[4]: 'Prof. Ray compares the author very properly to Tycho. But we should imagine him to be greater.' The importance of Samanta's work and the quality of his observations with naked eye have got true recognition through the comments made in the journal *Knowledge*[5]: 'Of all the numerous works on astronomy that have been published within the last few years, this is by far the most extraordinary and in some respects the most instructive. ....It is a complete system of astronomy founded upon naked-eye observations only. ....The work is of importance and interest to us Westerners also. It demonstrates the degree of accuracy which was possible in astronomical observation before the invention of telescope.' Here it will be probably fair to state that Samanta Chandra Sekhar belongs to the same class of Indian astronomers as Aryabhata, Varahamihira, Brahmagupta and Bhaskara.

It is unfortunate that Chandra Sekhar is almost unknown outside Orissa. This has so happened probably because he lived and worked during a period which was about 300 years after Orissa lost its pristine glory and preeminent position in the Indian subcontinent[13–16]. Although Orissa came under British rule in 1803, it continued to remain in virtual darkness during the lifetime of Samanta, being deprived of the advantage of modern education.

Almost a century has passed since the publication of *Siddhanta Darpana*. Its author passed away in 1904 living in the midst of poverty and struggle. The only recognition for the scientist during his lifetime was the acceptance of his astronomical prescriptions for regulation of the rituals at the Jagannath temple and the award of the title of Harichandan Mahapatra by Gajpati king of Puri in 1870. Of course, the title of Mahamahopadhyaya was conferred upon him by the British Government in 1893.

It is a matter of some relief that Kochhar and Narlikar[17] and SriRam[18] have recently referred to the work of Samanta Chandra Sekhar. Nevertheless, *Siddhanta Darpana* continues to serve the people of Orissa by providing the basis for the preparation of its almanac.

1. Simha Samanta, Mahamahopadhyaya Sri Chandra Sekhar Hari Chandan Mahapatra, *Siddhanta Darpana*, Indian Depository, Calcutta, 1899.
2. Mishra, Pandit Sri Chandrasekhar, *Samanta Chandrasekhar* (Biography in Oriya), Satyavadi Press, Cuttack, 1932.
3. Ray, J. C., *Introduction to Siddhanta Darpana*, 1899, pp. 1–56.
4. *Nature*, 1899, **59**, 436–437.
5. *Knowledge*, 1899, **XXII**, 257–258.
6. Simha Samanta, Mahamahopadhyaya Sri Chandrasekhar Simha Hari Chandan Mahapatra, *Siddhanta Darpana* (translated and described in Oriya by Pandit Sri Bira Hanuman Shastri; edited and compiled by Dayanidhi Khadiratna), Dharmagrantha Store, Cuttack, 1975.
7. Mahapatra Samanta, Mahamahopadhyaya Sri Chandrasekhar Simha Hari Chandan, *Siddhanta Darpana* (edited and translated to Oriya by Sri Bira Hanuman Shastri) Utkal University, Vanivihar, 1976.
8. Spencer, Jones Harold, *General Astronomy*, Edward Arnold, London, 1934, 2nd edn, chapt. VI, pp.114–117.
9. Barlow, C. W. C. and Bryan, G. H., *Elementary Mathematical Astronomy*, University Tutorial Press, London, 1944, 4th edn, pp.363–365.
10. Burgess, E. and Siddhanta, Surya (ed. Gangooly, P), Indological Book House, New Delhi, 1860, (reprinted 1977), p. 76.
11. Sengupta, P. C., *Khandakhadhyaka of Brahmagupta*, Calcutta University Press, Calcutta, 1934.
12. Ramsubramanian, K., Srinivas, M. D. and SriRam, M. S., *Curr. Sci.*, 1994, **66**, 784–790.
13. Nehru, J., *The Discovery of India*, Oxford University Press, 13th impression, 1993, pp. 132–133, 203–204.
14. Pannikar, K. M., *A Survey of Indian History*, Asia Publishing House, 1963, pp. 32, 47–48.
15. Banerjee, R. D., *History of Orissa*, Bharatiya Publishing House, New Delhi, 1980.
16. Mahatab, H. K., *The History of Orissa*, Cuttack Students' Store, 3rd edn., 1981.
17. Kochhar, R. and Narlikar, J., *Astronomy in India*, Inter-University Centre for Astronomy & Astrophysics, Pune, 1993, pp. 3, 42.
18. SriRam, M. S., *Man and the Universe*, Department of Theoretical Physics, University of Madras Manuscript, 1993, p. 199.
19. Lahiri, M. K., *Indian Ephemeresis of Planets' Positions for 1995 AD*, Astro-Research Bureau, Calcutta, 1994.
20. Arkasomayaji, D., *Siddhanta Siromani of Bhaskara*, Kendriya Sanskrit Vidyapitha, Tirupati, 1980.

ACKNOWLEDGEMENTS. The English introduction to *Siddhanta Darpana* by J. C. Ray and the Oriya Translation of the Treatise by Pandit Bira Hanuman Shastri have been our prime source of reference in preparation of this article. We are grateful to Dr Ranatosh Chakrabarti of the Department of Physiology of Surendranath College, Calcutta, for having been kind enough to provide us with the copies of the articles on Samanta Chandra Sekhar from *Nature* 1899 and *Knowledge* 1899. We are thankful to Dr S. Torasia for his cooperation and encouragement.

P. C. Naik is in the Pathani Samanta Planetarium, Bhubaneswar 751 013, India and L. Satpathy is in the Institute of Physics, Bhubaneswar 751 005, India.

Edited and published by Prof. S. Ramaseshan and Prof. P. Balaram, Current Science Association, Bangalore 560 080.
Typeset by Creative Literati Pvt. Ltd., Bangalore (Tel. cum Fax: 5462698), Printed at Printek, Bangalore (3357763)

INSTITUTE FOR PLASMA RESEARCH

(An Autonomous Institute of DST, Govt. of India)
Near Indira Bridge, Gandhinagar Highway,
BHAT, GANDHINAGAR 382 424

ADVT. NO. 08/95

FACULTY AND POST-DOCTORAL POSITIONS IN PLASMA PROCESSING

The Institute for Plasma Research is an autonomous research institute established by the Department of Science and Technology, Govt. of India to lead the country's Plasma Physics and Magnetic Confinement Fusion Research Programme.

The Plasma Processing Group of IPR is involved in the development of advanced plasma-assisted material processing technologies for the Indian industry. The group interacts with industry through all the phases of technology generation; from conception to commercialization. It is a multidisciplinary activity and hence provides considerable scope for exposure to different areas of specialization.

Plasma processing is an emergent technology which utilizes exotic properties of the plasma state to effect physical, chemical and metallurgical transformation in materials to produce new materials or impart new properties to conventional materials. Plasma processing has already entered the industrial arena in advanced countries.

The Plasma Processing Group of the Institute for Plasma Research exploits the IPR expertise plasma and allied technologies in developing plasma-assisted material processing technologies for Indian industries. PPG interacts closely with entrepreneurs through all the phases of technology generation and transfer from concept to commercialization. A technology demonstration facility and a surface and material characterization facility with many advanced instruments are being set up.

FACULTY AND POST-DOCTORAL POSITIONS

The PPG invites applications for the post of Associate Professor – I, Fellow/Scientist – SD and Post Doctoral Fellows. The specific projects on which they would be working are:

1. Development of processes and techniques in plasma-assisted pyrometallurgy and mineral processing.
2. Plasma-based synthesis of ceramic and metal powders of industrial relevance.
3. Development of multi-function surface engineering techniques and systems.
4. Plasma source ion-implantation techniques and material evaluation.

Qualification and Experience: Ph D with adequate research experience in high voltage techniques, surface sciences, ceramics, process and physical metallurgy and related areas to contribute to the research and development programmes.

GENERAL INFORMATION

Scale of Pay :

1. Associate Professor – I	:	Rs. 3700–125–4700–150–5000 (Gross Rs. 9730/-)
2. Fellow/Scientist – SD	:	Rs. 3000–100–3500–125–4500 (Gross Rs. 8080/-)
3. Post Doctoral Fellow	:	Consolidated Stipend Rs. 4000/- per month + Rs. 1000/- House Rent Subsidy.

Higher starting salary in the scale may be considered in exceptional cases.

Post nos. 1 and 2 are of regular nature and 3 is for one year tenure which may be extended for another term.

AGE LIMIT: 35 years in respect of Fellow/Scientist – SD and PDF. Relaxable by 5 years in case of SC/ST candidates.

The above norms are relaxable in highly exceptional cases. The Institute offers excellent career opportunities to creative, innovative and result-oriented staff. In addition to pay and allowances, perquisites like Contributory Provident Fund, Gratuity, Medical Facility under its Contributory Health Service Scheme, LTC, Group-Linked Insurance Scheme, Subsidised Transport and Canteen facilities are admissible as per Institute rules.

Out-station candidates called for interview will be paid first class to and fro rail/ordinary class bus fare by shortest route in respect of Associate Professor – I, Fellow/Scientist – SD and second class rail/ordinary class bus fare by shortest route in respect of PDF on production of proof of journey such as Railway/Bus ticket or ticket no.

How to Apply: Typed application on plain paper strictly in the following format with one recent passport size photograph duly affixed on it along with copies of marksheets, certificates, testimonials, etc. should reach Prof. P. I. John, Institute for Plasma Research at the above address [FAX: 91-079-7864310 E-MAIL: JOHN @ PLASMA.ERNET.IN] within 20 (twenty) days from the date of publication of this advertisement. **The envelope should be strictly superscribed with Advt. No. and Post Applied for.**

APPLICATION FORMAT

1. Advertisement No., 2. Post Applied for, 3. Name (in block letters), 4. Date of Birth, 5. Place of Birth, 6. Age, 7. Marital Status, 8. Present Address, 9. Permanent Address, 10. Whether belongs to SC/ST if so, attach documentary evidence, 11. Qualifications [SSC onwards, indicating (a) exam. passed, (b) board/university, (c) subjects, (d) year of passing, (e) division, (f) percentage of marks obtained], 12. Experience [indicating (a) name of the organization, (b) nature of work in detail, (c) period, (d) scale of pay and total salary drawn], 13. References, 14. Any other relevant information, 15. Signature with date.

Candidates employed in Govt./Quasi Govt./Public Sector undertakings should forward their applications through proper channel. The Institute reserves the right to reject any application without assigning any reason. **Applications not in the above format are liable to be rejected.** No interim correspondence will be entertained.

CURRENT SCIENCE

A fortnightly journal of research

Editors: S. Ramaseshan and P. Balaram

Editorial Board

S. Arunachalam, Central Electrochemical Research Institute, Karaikudi 623 006
Asis Datta, School of Life Sciences, Jawaharlal Nehru University, New Delhi 110 067
D. Balasubramanian, Centre for Cellular and Molecular Biology, Hyderabad 500 007
J. Chandrasekhar, Department of Organic Chemistry, Indian Institute of Science, Bangalore 560 012
R. Gadagkar, Centre for Ecological Sciences, Indian Institute of Science, Bangalore 560 012
K. N. Ganeshaiah, Department of Plant Genetics and Breeding, University of Agricultural Sciences, Bangalore 560 065
R. V. Hosur, Tata Institute of Fundamental Research, Bombay 400 005
N. V. Joshi, Centre for Ecological Sciences, Indian Institute of Science, Bangalore 560 012
Jyotirmoy Das, Indian Institute of Chemical Biology, Calcutta 700 032
C. C. Kartha, Division of Cellular and Molecular Cardiology, Sree Chitra Tirunal Institute for Medical Sciences and Technology, Thiruvananthapuram 695 011
S. S. Krishnamurthy, Inorganic and Physical Chemistry Department, Indian Institute of Science, Bangalore 560 012
S. C. Lakhotia, Department of Zoology, Banaras Hindu University, Varanasi 221 005
K. Muralidhar, Department of Zoology, Delhi University, Delhi 110 007
R. Narasimhan, CMC Limited, Bangalore 560 001
K. Neelakantan, ANURAG (Advanced Numerical Research and Analysis Group), Hyderabad 500 258
R. Nityananda, Raman Research Institute, Bangalore 560 080
T. Padmanabhan, Inter-University Centre for Astronomy and Astrophysics, Pune 411 007
B. R. Pai, Propulsion Division, National Aerospace Laboratories, Bangalore 560 017
T. J. Pandian, School of Biological Sciences, Madurai Kamaraj University, Madurai 625 021
Y. S. Rajan, Department of Science and Technology, New Delhi 110 016
M. Ramakrishnan, Geological Survey of India, Calcutta 700 016
A. V. Ramani, T. T. K. Pharma Ltd., Bangalore 560 025
S. Ramaswamy, Department of Physics, Indian Institute of Science, Bangalore 560 012
A. K. Raychaudhuri, Department of Physics, Indian Institute of Science, Bangalore 560 012
S. R. Shetye, Physical Oceanography Division, National Institute of Oceanography, Dona Paula, Goa 403 004
V. Siddhartha, 51 Bharati Nagar, New Delhi 110 003
V. Srinivas, School of Mathematics, Tata Institute of Fundamental Research, Bombay 400 005
G. Srinivasan, Raman Research Institute, Bangalore 560 080
R. Uma Shaanker, Department of Crop Physiology, University of Agricultural Sciences, Bangalore 560 065
K. S. Valdiya, Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore 560 064
G. Venkataraman, Department of Physics, Sri Sathya Sai Institute of Higher Learning, Prasanthi Nilayam 515 134
M. Vidyasagar, Centre for Artificial Intelligence and Robotics, Bangalore 560 001
K. VijayRaghavan, TIFR Centre, Indian Institute of Science Campus, Bangalore 560 012

Current Science, founded in 1932, is published by the Current Science Association in collaboration with the Indian Academy of Sciences. The journal is also intended as a medium for communication and discussion of important issues that concern science and scientific activity. All articles published in *Current Science*, especially editorials, opinions and commentaries, letters and book reviews, are deemed to reflect the individual views of the authors and not the official points of view, either of the Current Science Association or of the Indian Academy of Sciences.

Working Committee, Current Science Association

P. Balaram, Indian Institute of Science, Bangalore
R. Chidambaram, Department of Atomic Energy, Bombay
V. K. Gaur, CSIR Centre for Mathematical Modelling and Computer Simulation, Bangalore
K. Kasturirangan, ISRO Satellite Centre, Bangalore
P. K. Kaw, Institute for Plasma Research, Gandhinagar
N. Kumar, Raman Research Institute, Bangalore (*Vice-President*)
N. V. Madhusudana, Raman Research Institute, Bangalore (*Secretary*)
A. Mani, Field Research Unit, Indian Institute of Tropical Meteorology, Bangalore (*Past-President*)
R. A. Mashelkar, Council of Scientific and Industrial Research, New Delhi
R. Narasimha, Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore
G. Padmanaban, Indian Institute of Science, Bangalore (*Vice-President*)
R. Ramachandran, The Institute of Mathematical Sciences, Madras
P. Rama Rao, Defence Research and Development Organization, New Delhi
S. Ramaseshan, Raman Research Institute, Bangalore (*Treasurer*)
M. S. Valiathan, Manipal Academy of Higher Education, Manipal (*President*)

Editorial Office

Editorial staff

Chandrika Ramesh, G. Madhavan, T. D. Mahabaleswara, C. S. Ravi Kumar, M. S. Venugopal

Circulation and Accounts

Peter Jayaraj, Ranjini Mohan, B. Sethumani, Shanthi Bhaskar, B. K. Shivaramiah

Annual subscription 1995

	Institutional	Personal
India	Rs 350	Rs 100
All other countries (AIR MAIL)	US $175	US $50

Single copies other than special issues: Rs 35/US $15

Editorial Office: CURRENT SCIENCE
C. V. Raman Avenue
P.B. 8001
Bangalore 560 080, India
Telephone: 91-80-3342310
Telex 845-2178 ACAD IN
Fax: 91-80-3346094

Copyright © 1995, Indian Academy of Sciences

INFORMATION FOR CONTRIBUTORS

GENERAL

All manuscripts should be addressed to the Editor, *Current Science*, P. B. No. 8001, C. V. Raman Avenue, Bangalore 560 080. Submission of an article will be held to imply that it has not been previously published and is not under consideration for publication elsewhere; and further, that if accepted, it will not be published elsewhere. *Three copies of contributions of all categories* are required, with a letter of transmittal giving (i) names and complete addresses of the authors and (ii) title of the contribution and the category in which it is submitted (see below).

Current Science is a multidisciplinary journal and therefore research and review papers of general significance that are written clearly and well organized will be given preference. All papers will be first assessed by a Reviewing Editor. Papers found unsuitable in terms of the overall requirements of the journal will be returned to the authors. The others will be sent for detailed review. *Both solicited and unsolicited material will be reviewed.* Authors of these papers will be notified of acceptance, rejection, or need for revision of the paper. Returned papers cannot be resubmitted.

Illustrations and other material to be reproduced from other publications must be properly credited; it is the authors' responsibility to obtain permission for reproduction (copies of letters of permission should be sent).

CATEGORIES OF MANUSCRIPT

General articles (not exceeding 5000 words) discuss current trends in research in a field that will be of interest to readers outside the field; interdisciplinary topics: science policy and administration; or some aspect of the application of science and technology to human needs or the impact of science and technology on society/ecosystems/life. They should include a summary not exceeding 100 words, introductory paragraph(s), brief subheads at appropriate places to point to what follows, illustrations that will help a general reader, and references.

Review articles (not exceeding 5000 words) are expected to survey and discuss recent developments in a field. They should be well focused and organized, and avoid a general, 'textbook' style.

Research articles (not exceeding 4000 words) should report research results of fairly major significance. They should include an abstract not exceeding 100 words, introductory paragraph(s), and brief subheads.

Research communications (not exceeding 2000 words) should contain important findings that are novel and of fairly broad interest. They should include a brief abstract and an introductory paragraph. Text should not be broken up under subheads.

Correspondence includes letters, not exceeding 500 words, that are of general interest to scientists. All letters cannot be published.

Scientific correspondence contains technical comments, including those on articles or communications published in *Current Science* within the previous six months. Letters may be reviewed and edited.

Research news articles are intended to inform nonspecialists about recently published advances or important findings discussed at a meeting. Authors should also send a copy of the paper(s) on which the article is based. Meeting reports should avoid merely listing brief accounts of topics discussed, and must convey to readers the significance of an important advance.

Research accounts articles are intended to be personalized reviews of research from the authors' own laboratory, based on a body of published work. The articles must provide appropriate background to the area in a concise introduction, which should also serve to place the author's work in proper perspective. Articles will normally not exceed 8 to 10 printed pages.

Opinion articles present views on issues related to science and scientific activity. **Commentary** articles should contain expository notes on issues related to science and scientific activity.

Book reviews. Unsolicited reviews will also be considered. Reviews that merely 'list' brief descriptions of the contents cannot be published. Reviews should have 'context' and convey some information about the subject of the book.

Historical commentary and notes inform readers about interesting aspects of personalities or institutions of science or about watershed events in the history/development of science; most are expected to relate to India. Illustrations are welcome. Brief items will also be considered.

MANUSCRIPT PREPARATION

Manuscripts should be typed double-spaced on one side of white bond paper (21 × 28 cm). The pages should be numbered consecutively, starting with the title page and through the text, reference list, tables and figure legends. The **title** should be brief, specific and amenable to indexing. Not more than five **keywords** should be indicated separately; these should be chosen carefully and must not be phrases of several words. **Summary** and **abstract** should not have more than 100 words and should convey the main point of the paper, outline the results and conclusions, and explain the significance of the results.

Text. All papers should have a brief introduction. The text should be intelligible to readers in different disciplines and technical terms should be defined. Tables and figures should be referred to in numerical order. All **symbols** and **abbreviations** must be defined, and used only when absolutely necessary. Superscripts and subscripts and ambiguous characters should be clearly indicated. **Units of measure** should be metric or, preferably, SI. Methods should, as far as possible, be described briefly in appropriate table and figure legends.

Figures. In the case of line drawings one set of originals (without any lettering) is sufficient, with two copies containing lettering. In the case of photographs good prints are required with each copy of the manuscript; photocopies are not acceptable. Line drawings should be roughly twice the final printed size. The correct orientation should be indicated if not clear.

Photomicrographs and other photographs that require it must have a scale bar, which should be defined clearly in the legend. Primary data should be submitted as far as possible (e.g. actual photographs of electrophoretic gels rather than idealized diagrams).

References should be numbered in the order in which they appear, first through the text and then through table and figure legends. The following are examples of ways of writing references.

1. Mukundan, T. and Kishore, K., *Curr. Sci.*, 1991, **60**, 355–362.
2. Constantine. G., in *Biology of Bats* (ed. Wimsatt. W. A), Academic Press, New York. 1970. vol 1, pp. 319–322.

Acknowledgements should be brief. Footnotes are not allowed except to identify the corresponding author if not the first.

Cover photographs. Good photographs (colour or black and white) that pertain to a submitted paper will be considered for use on the cover. Good *prints* and a legend should be submitted with the manuscript. In the case of a colour picture, a transparency will be required for printing if accepted.

PROOFS AND PUBLICATION

Two sets of galley proofs are sent to the corresponding author. A reprint order form accompanies the proofs.

Founded: 1932

ISSN: 0011-3891

CURRENT SCIENCE

Volume 69 Number 9

10 November 1995

715 In this issue

CORRESPONDENCE

717 **Microbiology in India**
P. Tauro

717 **Feast of the Gods?**
R. R. Upadhyay

NEWS

717 **Nobel Peace Prize for 1995 to Joseph Rotblat and the Pugwash Conferences on Science and World Affairs**

SCIENTIFIC CORRESPONDENCE

718 **A note on the occurrence of Columbite–Tantalite-bearing pegmatite at Limboi, District Sabarkantha, Gujarat**
N. Muralikrishna, K. K. Pandey, P. K. Shrivastava, S. Ramachandran and Rajendra Singh

RESEARCH NEWS

719 **Discovery of the new radiation belt of the earth**
Sukumar Biswas

GENERAL ARTICLE

721 **Gateways to the galaxy: New findings in cosmic rays from Spacelab-3 experiment 'Anuradha'**
Sukumar Biswas and Jitendra Nath Goswami

REVIEW ARTICLES

732 **Cosmic gamma ray bursts – Recent developments and observations from SROSS satellites**
K. Kasturirangan, V. N. Padmini, N. L. Prasad and U. R. Rao

739 **Quantum Monte Carlo techniques: Chemical applications**
Charusita Chakravarty

747 **Vitamin C: A potential saviour against free radical-induced oxidative damage**
Indu B. Chatterjee, Chinmay K. Mukhopadhyay and Mrinal K. Ghosh

RESEARCH ARTICLE

752 **Effect of transient increase of greenhouse gases and sulphate aerosols on monsoon climate**
M. Lal, U. Cubasch, R. Voss and J. Waszkewitz

RESEARCH COMMUNICATIONS

763 **Surinamite (Be, B, Ga) from cordierite gneisses of eastern ghat mobile belt, India**
P. V. Ramesh Kumar, K. K. V. S. Raju and B. S. Ganga Rao

767 **Evidence of garnet to spinel peridotite transition in the harzburgites of Indus ophiolite belt: An indication for their mantle origin**
N. R. Karmalkar, A. G. Dessai, R. A. Duraiswami and M. Leblanc

771 **Bilateral outgrowths on thorax of sepia eye colour mutants in *Drosophila bipectinata***
B. N. Singh, Seema Sisodia and Rakhee Banerjee

772 **A new empirical relation for strong seismic ground motion for the Himalayan region**
Ashutosh Aman, Umesh K. Singh and Ramesh P. Singh

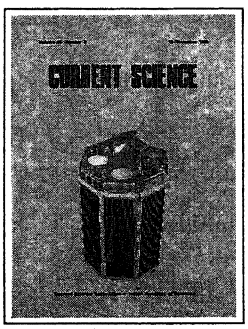
777 **Involvement of calcium in brassinolide and auxin-induced cell elongation**
Anitha Naren, T. G. Prasad, V. R. Sashidhar and M. Udaya Kumar

BOOK REVIEW

781 **Uttarkashi Earthquake,** *reviewed by* Ramesh Chander

GENERALIA

783 **In the pursuit of science: The 'subjective' element**
A. N. Mitra



COVER. Photograph of the SROSS-C satellite. See page 732.

Indexed in CURRENT CONTENTS/GEOBASE/CHEMICAL ABSTRACTS

In this issue

Investigating cosmic rays and gamma ray bursts

The present century has witnessed many revolutions in astronomical techniques. Gone are the days when all the information about the Universe used to be collected through a narrow band of the electromagnetic radiation spectrum, for which we are endowed with a natural detector—the human eye. First to appear outside this restricted channel was cosmic rays, discovered by Hesse in 1912, followed by radio and infrared, and then higher energy radiations with the opening up of the space window. In this issue we have two excellent articles reviewing two of such new techniques, giving first hand knowledge from their own scientific efforts.

'Cosmic rays', named by Millikan shortly after the discovery, is not a happy choice, as the pair of words conceal two misconceptions about the phenomena: These are not rays, but streams of very high energy particles, and we are still at a loss to understand their origin. To confound the issue further 'primary cosmic ray' particles are not directly detectable from ground being extensively modified by our atmospheric envelope. So, efforts were made over decades to hoist instruments up by balloons, rockets, satellites and space probes beyond our near space environment. In this direction the Indian collaborative effort 'Anuradha' has been a remarkable success; the story of its conception, design, fabrication and flight in a space shuttle has been described (**page 721**) by the leader of the project team, unfolding some glimpses of new findings of the nature of these particles.

The second topic covers the very high energy radiations from some cosmic sources; the article has been written by a team of scientists from ISRO who themselves have made considerable contributions. Gamma ray bursts were serendipitously discovered from some satellite experiments in late sixties; the observations had posed a puzzle, and ever since efforts towards unfolding their secrets are being pursued. The Indian scientific team has designed, fabricated new instruments and have them installed in two of their multi-purpose satellites. The article (**page 732**) describes the mystery surrounding the phenomena and suggests new efforts needed for understanding them.

J. C. Bhattacharyya

Going beyond the Born–Oppenheimer description

One of the greatest challenges to chemical theory is the derivation of molecular level understanding of macroscopic properties. Two types of computational methods are currently employed for this purpose. First, single molecules or small clusters of them are studied using quantum chemical methods. Although the systems consist of interacting nuclei and electrons, a common approach is to treat the problem in stages. For a given frozen nuclear configuration, the electronic energy is computed. The process is repeated for different geometries to obtain a multi-dimensional potential energy surface. The problem of what the nuclei do in such a surface is solved subsequently. When a large collection of molecules is involved, Monte Carlo and Molecular Dynamics represent the best methods to use. The range of geometric configurations the system can adopt is generated by a Boltzmann sampling or by solving Newton's equation of motion. Of course, these calculations often use empirical potential functions and do not necessarily rely on PE surfaces derived from quantum chemical calculations.

An alternative computational strategy is becoming increasingly popular. It is to treat nuclear and electronic motions at the same time. The energetics of a system as it dynamically responds to changes in nuclear and electronic coordinates is evaluated using Quantum Monte Carlo methods. Technical details as well as computational nuances of these procedures are reviewed by C. Chakravarty (**page 739**). Some representative applications of the methods to interesting chemical problems are also discussed.

J. Chandrasekhar

Fighting oxidative damage with vitamin C

Linus Pauling's enthusiastic espousal of the use of mega doses of vitamin C for staving off the common cold and cancer may not have found much favour in the scientific establishment. The Pauling campaign however promoted the widespread use of vitamin C and millions around the world probably have their own tales to tell of the benefits of regular consumption. The vitamin itself was first discovered accidentally, with the discoverer, Albert Szent-Györgi having no pretensions to belonging to the heroic tribe of vitamin hunters of the early decades of this century. Indeed, Szent-Györgi irreverently christened the white crystals, 'ignose' ('ig' for ignorance and 'ose' for sugar). Hexuronic acid and ascorbic acid were names that came later. While vitamin C was originally identified as the anti-scurvy factor its biochemical functions are diverse, stemming almost exclusively from its anti-oxidant properties. Oxidative processes are associated with many degenerative conditions and aging; unsurprisingly antioxidants are frequently touted as the universal panacea against a wide spectrum of physiological disorders.

In India, research on the biosynthesis of ascorbic acid was pioneered in the 1950s by B. C. Guha in Calcutta. This tradition has been well maintained in many of the researches on vitamin C that are described by I. B. Chatterjee and coworkers on **page 747**.

P. Balaram

One Day Symposium on GENOME ANALYSIS: RECENT TRENDS AND APPLICATIONS

16 January, 1996

Invited speakers

A Bhattacharya	JNU, New Delhi
S K Brahmachari	I I Sc., Bangalore
G Padmanaban	I I Sc., Bangalore
J M Robertson	Applied Biosystems, USA
S K Sen	Bose Institute, Calcutta
E M Southern	Oxford University, UK

(Symposium will be concludeded by Panel Discussion on legal and ethical issues arising out of Human Genome Analysis)

Registration Fee Rs. 300/-

Last date of registration
31 December, 1995

Hands-on Training Course on DNA FINGERPRINTING

17 January - 4 February, 1996

* Course covers topics such as bacterial, human and population genetics, structure and sequencing of DNA, principles and applications of DNA fingerprinting, detection of genetic disorders by RFLP, fluorescence *in situ* hybridization (FISH), etc.

* Course also offers hands-on training in isolation of plasmid and bacterial DNA, vector-host systems, molecular cloning, *FISH*, DNA fingerprinting of samples of human/animal/plant origin.

Registration Fee Rs. 20,000/-
(inclusive of registration fee for symposium)

Last date for submission of application
30 November, 1995

Organized by

Centre for Cellular and Molecular Biology (CCMB), Hyderabad 500 007, INDIA

Sponsored by

Centre for DNA Fingerprinting and Diagnostics (CDFD), Hyderabad 500 007

Association for the Promotion of DNA Fingerprinting and Other DNA Technologies (APDFDT) Hyderabad 500 007

Department of Biotechnology (DBT), Government of India New Delhi 110 003

Council of Scientific and Industrial Research (CSIR) New Delhi 110 001

For further details contact

Dr Lalji Singh
Course Director
Centre for Cellular and Molecular Biology
Uppal Road, Hyderabad 500 007

PHONE: 040 671252
TELEX: 0425 7046 CCMB IN
FAX: 040 671195
E. Mail: lalji@ccmb.uunet.in

CURRENT SCIENCE

Volume 69 Number 9 10 November 1995

CORRESPONDENCE

Microbiology in India

Ramesh Maheswari's article on 'Whither microbiology? (*Curr. Sci.*, 1995, **69**, 401–406) should be an eye opener to all those microbiologists in India who think that salvation exists only in gene cloning and genetic engineering. We have not even seen a tenth of the microbial world and there is a lot more to explore and understand. What we have understood so far should form the basis of future biology, as Maheswari points out. The lopsided support that microbiology receives is one reason for believing that microbiology is not dead but is moribund. Neglect of classical microbiology is the reason why our national committee on plague has so far been unable to say what exactly caused the Surat epidemic. Carl Woese's thesis is only partly correct. Such ups and downs are seen in every branch of science. The reemergence of some of the well-known diseases is one instance where the Western world has realized the importance of classical microbiology. Let us not be discouraged by all that people in the West have to say. After all, all that comes from the West need not be the best. Undoubtedly, there is a need to reorient our teaching and research in microbiology if we are to meet the challenges of the 21st century. It is a pity that although this subject has been taught in more than 40 Indian universities under the UGC cover both at undergraduate and postgraduate levels, there is as yet no attempt made to evolve a standard undergraduate curriculum. There is no need to despair but there is certainly a need to evolve a new trend in teaching microbiology to include both classical and modern aspects.

P. TAURO

Department of Microbiology,
Haryana Agricultural University,
Hisar, India

Feast of the Gods?

Total solar eclipse – October 1995 (J. C. Bhattacharyya, *Curr. Sci.*, 1995, **69**, 486–488), will indeed be a spectacular event, which is bound to be fascinating for astrophysicists and equally for enthusiastic viewers. The mention of the comments of Dr M. K. V. Bappu, on the total solar eclipse of 1980, is full of aesthetic sense of feelings but the attempt to mix science with the feast of the Gods or for that matter with God even, sounds a strange contradiction.

R. R. UPADHYAY

Upadhyay Cancer Research Institute,
Campus Kothi Kake Babu,
Deo Kali Road,
Faizabad 224 001, India

NEWS

Nobel Peace Prize for 1995 to Joseph Rotblat and the Pugwash Conferences on Science and World Affairs

The Norwegian Nobel Committee has awarded the Nobel Peace Prize for 1995, in two equal parts, to *Joseph Rotblat* and to the *Pugwash Conferences on Science and World Affairs*, for their efforts to diminish the part played by nuclear arms in international politics and in the longer run to eliminate such arms.

It is fifty years this year since the two atomic bombs were dropped on Hiroshima and Nagasaki, and forty years since the issuing of the Russell–Einstein Manifesto. The Manifesto laid the foundations for the Pugwash Conferences, which have maintained a high level of activity to this day. Joseph Rotblat was one of the eleven scientists behind the Manifesto, and has since been the most important figure in the Pugwash work.

The Conferences are based on the recognition of the responsibility of scientists for their inventions. They have underlined the catastrophic consequences of the use of the new weapons. They have brought together scientists and decision-makers to collaborate across political divides on constructive proposals for reducing the nuclear threat.

The Pugwash Conferences are founded in the desire to see all nuclear arms destroyed and, ultimately, in a vision of other solutions to international disputes than war. The Pugwash Conference in Hiroshima in July this year declared that we have the opportunity today of approaching those goals. It is the Committee's hope that the award of the Nobel Peace Prize for 1995 to Rotblat and to Pugwash will encourage world leaders to intensify their efforts to rid the world of nuclear weapons.

A note on the occurrence of columbite–tantalite bearing pegmatite at Limboi, District Sabarkantha, Gujarat

Although some rare earth and rare metal minerals like gadolinite ($Be_2 FeY_2 Si_2O_{10}$) associated with distinct large crystals of tin, from some tourmaline pegmatite of Palanpur–Hosainpur area[1] and pyrochlore ($Na_2CaNb_2O_6$) from carbonatite of Ambadongar area[2] are recorded in Gujarat, the presently reported columbite–tantalite (Fe, Mn) O (Nb, Ta)$_2O_5$ mineralization, which occurs as a discrete mineral phase in the granitic pegmatite in the Limboi area, district Sabarkantha, is first of its kind located in Gujarat.

Columbite–tantalite mineralization in the deluvial/eluvial pegmatite gravel around the quartz core, with a grade of 0.03–0.06% of the mineral by weight, has been established by preliminary test-pitting and panning at Limboi. Columbite–tantalite pieces vary from 5 to 50 mm across and a sample analysed by X-ray fluoroscence (XRF) has 60.07% Nb_2O_5, 15.35% Ta_2O_5, 1.29% Y_2O_3, 1559 ppm Sn, 0.30% Ce_2O_3 and less than 20 ppm Ba. Columbite–tantalite (confirmed by X-ray diffraction studies) is radioactive and has assayed 1.2% eU_3O_8, 0.2% U_3O_8 and 2.4% ThO_2, by conventional radiometric methods. Chemical analysis of another columbite–tantalite sample from the Limboi occurrence has shown 53.6% Nb_2O_5, 12.1% Ta_2O_5, 1.5% TiO_2, 2.71% Y_2O_3, 1.3% WO_3, less than 0.1% SnO_2 besides 0.288% U_3O_8 and 1100 ppm ThO_2. Other ore minerals that are present at Limboi include fersmite (Ca, Ce, Na) (Nb, Ta, Ti)$_2$ (O, OH, F)$_6$, fluorite and beryl.

The Limboi columbite–tantalite pegmatite, emplaced within the Idar granite pluton is an interior type of zoned granitic pegmatite which trends NNE–SSW and measures 350 × 200 m (Figure 1) and forms two distinct mounds. It comprises a prominent and highly fractured quartz core surrounded by an intergrowth zone consisting of quartz, K-feldspar, muscovite and tourmaline with columbite–tantalite, fersmite, fluorite and beryl. This zone is largely covered by soil and eluvial gravels. High yttrium content up to 2.71% Y_2O_3 in the columbite–tantalite samples from this pegmatite may suggest possible presence of xenotime, in the vicinity.

In view of a close cluster of pegmatites and massive quartz veins in the Idar–Verabar ghat section, the area in the neighbourhood of Limboi pegmatitic occurrence was further examined, to explore the possibilities of locating some more columbite–tantalite occurrences. This resulted in the location of three more pegmatite occurrences, viz. Sarangpur, Sabalwada and Bhavangadh (Figure 2). Columbite–tantalite samples from Sarangpur (23°51′47″N : 72°57′41″E, Toposheet No. 46A/13) have shown 51.80 to 52.89% Nb_2O_5, 21.07 to 21.37% Ta_2O_5 and 1.87 to 2.01% SnO_2.

The proterozoic rocks of the north Gujarat comprise mainly the metasedimentary sequences of the Aravalli and Delhi Supergroups with associated synkinematic as well as post kinematic, (Post Delhi), late Proterozoic granites. Post tectonic suites of intrusive granitic rock with dominant pegmatitic phases characteristically mark the close of igneous activity. The intrusion of granitoids during the period 950–700 Ma also resulted in widespread migmatization of the preexisting rocks in North Gujarat and SW Rajasthan, in the southern part of Aravalli–Delhi fold belt. Such younger phases of granitoids are known to yield residual fluids, which could be anomalously rich in rare metals and rare earths, depending upon the degree of fractionation. The trace element geochemistry and geochronology of the Idar granite[3] clearly correspond with the Erinpura granitic suite of rocks, dated 740 ± 10 Ma (ref. 4), occupying large tracts in the adjoining Banaskantha district (Gujarat) and Sirohi district (Rajasthan). These are highly differentiated granites and constitute younger fertile phases which might have contributed rare metals and rare earths to the mid-Proterozoic metasediments of Delhi Supergroup, in the contact zones and peripheral parts of the granite batholith.

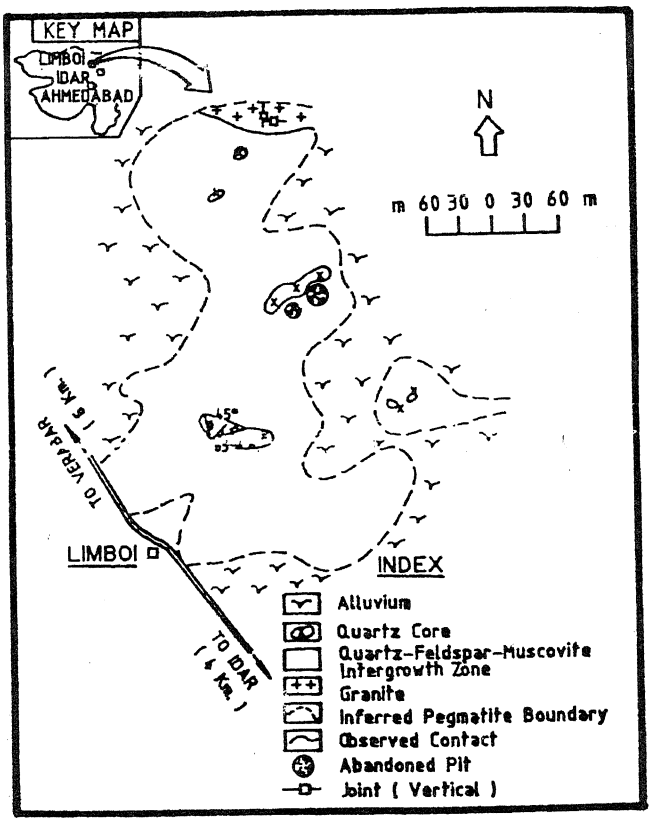


Figure 1. Geological map of Limboi pegmatite, Sabarkantha district, Gujarat.

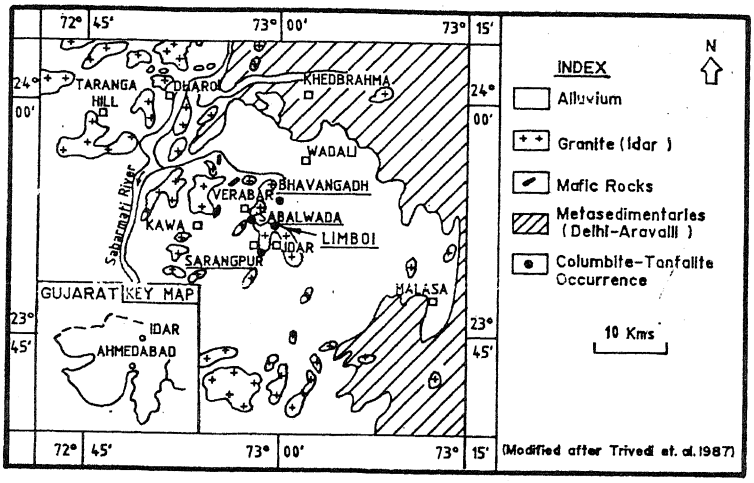


Figure 2. Geological map of Idar Area, North Gujarat, showing columbite–tantalite occurrences.

The presence of columbite–tantalite-bearing pegmatites around Limboi and granitic terrain occupying about 1200 km^2 have opened up vast area for exploration of rare metals and rare earths. Follow-up investigations in the area around Verabar, Kawa, Dharoi, Nadri, Nawawas, in the Sabarkantha district of Gujarat and also in the adjoining Banaskantha district of Gujarat and Sirohi district of southwestern Rajasthan are likely to be rewarding, as fertile granites are present in the area.

1. Clegg, E. L. G., *Rec. Geol. Surv. India*, 1944, 76, 107.
2. Udas, G. R., *Geol. Surv. India*, Special publication No. 13, 1983, pp. 37–40.
3. Trivedi, J. R., Gopalan, K. and Patel, P. P., in *Recent Researches in Geology* (ed. Saha A. K.), Hindustan Publishing Corporation, 1987, vol. 13, pp. 77–83.
4. Shrivastava, P. K., Khandelwal, M. K., Despati, T., Khirwadkar, M. P. and Dwivedy, K. K., *Exploration and Research for Atomic Minerals*, 1990, vol. 3, pp. 45–56.

ACKNOWLEDGEMENT. We gratefully acknowledge the support, guidance and encouragement received from K. K. Dwivedy, Director, AMD, Hyderabad during the course of field investigation and also for granting permission to publish this note.

N. MURALIKRISHNA
K. K. PANDEY
P. K. SHRIVASTAVA
S. RAMACHANDRAN
RAJENDRA SINGH*

Atomic Minerals Division,
Western Region,
R. V. Desai Road,
Baroda 390 001, India
**Atomic Minerals Division,*
Begumpet, Hyderabad 500 016, India

RESEARCH NEWS

Discovery of the new radiation belt of the earth

Sukumar Biswas

The discovery of the new – the third-radiation belt of the earth has followed from the results obtained with the help of earth-orbiting satellites[1]. This discovery was made in three stages. The first evidence came in a striking manner in the Indian cosmic-ray experiment[2,3] in the SKYLAB-III mission of NASA during 1973–74. Further evidence was obtained from the results of the COSMOS satellites[4] of USSR flown in 1985–88. Finally, the confirmation and identification of the new belt was made from the results of the SAMPEX satellite[5] of NASA of USA in 1993.

Firstly, let us have a brief look at the two well-known radiation belts of the earth, the inner and outer belts discovered in 1958–59. These are also called Van Allen Belts after their discoverer, J. A. Van Allen of USA. A number of satellite experiments were needed to understand their properties and origin. The inner belt consists of energetic charged particles, mostly protons (of energy ~ 10 to 100 MeV) and some electrons which are confined in a dough-nut shaped volume around the earth by the earth's magnetic field. The charged particles travel in spiralling paths around the magnetic field lines and they oscillate between the north and south hemispheres. At the same time they circle around the earth, forming a belt-like structure. The centre of the inner belt is located at about 3,600 km altitude where intensity of fast protons is enormously high. Similarly, the outer belt, which is composed of mostly fast-moving electrons, extends from ~ 10,000 to 70,000 km altitude. It has extremely high intensity of fast electrons near its core at 20,000–30,000 km. These trapped radiation belts present serious radiation hazards to man and instruments in these regions of space.

A new kind of cosmic rays were first seen in interplanetary space in experiments on Pioneer 10 and 11 and IMP-8 spacecrafts of USA in 1973–74 (refs. 6–8). These new components of cosmic rays, called anomalous cosmic rays, were first seen in near-earth space, inside the earth's magnetic field in the Indian cosmic-ray experiment[2,3] on the Skylab-III mission of NASA. The surprising result found in this experiment was that the measured intensity in the solid state track detectors in Skylab was at least 25 times higher than the calculated value from the data in interplanetary space. How could it be possible?

It was proposed by the Indian group[3] that some of the anomalous cosmic-ray

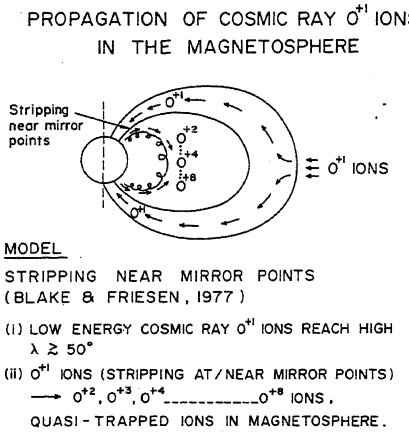


Figure 1. A sketch of the model of the origin of the newly-discovered radiation belt of the earth originating from the trapped anomalous cosmic ray particles, such as O^{1+} (from ref. 3). For details see text.

particles were trapped in the earth's magnetic field and these trapped or quasi-trapped particles were responsible for the excess flux. The mechanism for trapping was also suggested[3] by the Indian group. This mechanism, which is discussed a little later, is now accepted as the model for the origin of the trapped anomalous cosmic rays. In the Skylab experiment it was, however, not possible to locate the position of the new trapped component, for which more detailed experimentation was necessary.

A few attempts were made in early 1980s on this problem, but without success. In 1985–88, the USSR group made several COSMOS satellite flights with solid state track detectors, in some of which they noted double-peaked angular distributions of the oxygen ions of anomalous cosmic rays[4]. These were interpreted as arising due to the presence of a trapped component. Thus the evidence for the trapped component became stronger.

In 1992–94, the Caltech and NASA groups of USA flew a sophisticated mass spectrometer experiment in a NASA satellite, SAMPEX, whose orbit was at a higher altitude of 670 km. It was a polar-orbiting satellite with orbital inclination of 82° to the equator. The results from this experiment yielded the new evidence of the trapped anomalous cosmic rays[5], which confirmed earlier observations. It showed a distinct component which was due to trapped heavy ions of anomalous cosmic rays (ACR) and was composed mainly of oxygen ions of about 10–50 MeV/N. They located the new radiation belt of trapped ACR at a distance of about 6400 km above the equator. The trapped particles are confined in a torus-like region around the earth's equatorial region. From this volume a small fraction of trapped particles leaks out in high altitude regions of the upper atmosphere above the South Atlantic ocean region and these were detected by the earth-orbiting satellites when they passed through this region.

The mechanism for the formation of the trapped component of anomalous cosmic rays was proposed by the Indian group[3] and this is shown schematically in Figure 1, which is reproduced from the paper of Biswas and Durgaprasad[3]. In this model it is proposed that singly ionized ions of anomalous cosmic rays, such as O^{1+}, enter the geomagnetic field at high latitudes, say $\geq 50°$. Here a fraction of these suffer electron stripping process in the upper atmosphere near their mirror points according to Blake–Freisen mechanism[9]. The stripped ions having lower magnetic rigidities (i.e. momentum/effective charge) are trapped, forming the trapped component of anomalous cosmic rays. One may note that in this picture the ACR ions should be singly ionized; several experiments have determined that ACR ions of O, N, Ne are singly ionized[10]. The trapped heavy ions of ACR are stored in the belt for periods ranging from less than a day to about one year, depending on the ionization loss in the thin rarified regions of space where gas densities range from about 10^3 to 10^6 atoms/cm^3.

One may ask about the scientific usefulness or otherwise of this new radiation belt. The new belt is of weak intensity and so does not present any additional radiation hazard compared to the high intensity Van Allen belts. This belt, however, provides a new advantage. It is now established that ACR particles originate from neutral atoms of nearby interstellar space[10]; therefore the new belt of trapped ACR particles brings to us a sample of interstellar matter in near-earth space. Thus new properties of interstellar neutral matter, e.g. their isotopic composition, can be studied and new, important astrophysical knowledge may be obtained.

1. Biswas, S., *Space Sci. Rev.*, 1995, in press.
2. Biswas, S. *et al.*, *Astrophys. Space Sci.*, 1975, **35**, 337–347.
3. Biswas, S. and Durgaprasad, N., *Space Sci. Rev.*, 1980, **25**, 285–327.
4. Grigorov, N. L. *et al.*, Proceedings of the 21st International Cosmic Ray Conference, Adelaide, 1990, vol. 6, pp. 176–179.
5. Cummings, J. R. *et al.*, Proceedings of the 23rd International Cosmic Ray Conference, Calgary, 1993, vol. 3, pp. 428–431.
6. Mc-Donald, F. B. *et al.*, *Astrophys. J.*, 1974, **187**, L105.
7. Hovestadt, D. *et al.*, *Phys. Rev. Lett.*, 1973, **31**, 650.
8. Garcia-Munoz, M. *et al.*, *Astrophys. J.*, 1973, **182**, L81.
9. Blake, J. B. and Freisen, L. M., Proceedings of the 15th International Cosmic Ray Conference, Plovdiv, 1977, vol. 2, pp. 341–344.
10. Biswas, S. *et al.*, *Space Sci. Rev.*, 1993, **62**, 3–65.

Sukumar Biswas is in the Cosmic Ray Space Physics Group, Tata Institute of Fundamental Research, Bombay 400 005, India.

Gateways to the galaxy: New findings in cosmic rays from Spacelab-3 experiment 'Anuradha'

Sukumar Biswas and Jitendra Nath Goswami

The Anuradha cosmic-ray experiment, flown on board Space Shuttle Spacelab-3, remains one of the most successful Indian space science experiments conducted so far. This experiment provided the first direct and unambiguous results on the ionization states of the newly discovered anomalous cosmic rays that conclusively established the interstellar neutral atoms as their source. New and unexpected results on the properties and composition of low-energy cosmic ray heavy nuclei were also obtained. The scientific and technical activities that made this experiment a reality—from the planning stage to its operation during the Spacelab-3 mission, and the details of the post-flight data analysis and synthesis that provided the new results are described. Some new possibilities have emerged for future space experiments in this area.

Cosmic rays: A sample of galactic matter

LIKE many branches of science cosmic rays had a humble beginning. In the early part of this century, soon after the discovery of radioactivity, scientists were puzzled to note the leakage of electric charge from well-insulated bodies. To explore this problem, the German scientist H. Hess had flown well-insulated electroscopes in balloons in 1912, in one of which he himself was a passenger, and discovered that a penetrating radiation is impinging on the earth from extra-terrestrial sources. It was named cosmic rays by the American scientist R. A. Millikan, as the Sun could be ruled out as a source of this radiation. The energies of cosmic-ray particles are really astounding—it covers the vast span of about 10^6 eV to as high as 10^{20} eV. With instruments flown in high-altitude balloons it was discovered that cosmic-ray particles are composed of atomic nuclei of various elements such as hydrogen, helium, carbon, nitrogen, oxygen, etc. up to iron. In fact elements up to uranium are also present in cosmic rays with small intensities. The cosmic rays thus provide us with an unique sample of matter from interstellar space and opened to us a new window to the universe—to explore high-energy processes in the galaxy and beyond. On anticipating the fascinating discoveries in the field of cosmic rays, the eminent scientists, J. J. Thomson and J. P. Thomson remarked in 1928: 'It would be one of the romances of science if these obscure and prosaic minute leakages of electricity from well-insulated bodies would be the means by which the most fundamental problems in the evolution of the cosmos came to be investigated.' In the past decades many studies were made in spacecrafts to unravel the mysteries of the origin and acceleration of cosmic rays to high energies. In spite of spectacular progress, many of these problems still remain obscure and elusive. At present supernovae, their remnants, and pulsars are considered as likely sources of galactic cosmic rays; some of the radio-galaxies, quasars and active galactic nuclei could be possible sources of extragalactic cosmic rays. These studies have captured the fascination of numerous scientists in the world as high energy cosmic gateways to the galaxy and to the universe[1].

Sukumar Biswas is in the Tata Institute of Fundamental Research, Bombay 400 005, India, and Jitendra Nath Goswami is in the Physical Research Laboratory, Ahmedabad 380 009, India.

The discovery of anomalous cosmic rays

In the early 1970s, a new component of cosmic rays was discovered by analysing the data recorded in cosmic-ray detectors in Pioneer 10 and 11 spacecrafts that were on their way to Jupiter and in the interplanetary space probe IMP-8 orbiting the earth[2–4]. This component of low-energy cosmic rays of extra solar origin has some surprisingly new properties. In galactic cosmic rays, we normally have about 90% proton and ~ 9% of He, while C, N, O and heavier nuclei constitute ~ 1%. In contrast, in the new component of low-energy cosmic rays, protons were not detected, abundances of helium, oxygen, nitrogen and neon were enhanced, while other elements such as carbon, magnesium, silicon and iron had very low abundance or almost absent. As for

example, in the new component the abundance ratio of O/C is ≥ 5 in contrast to O/C ~ 1 in galactic cosmic rays (GCR). Because of its unusual composition and energy spectra, this new component of low-energy cosmic rays was called the Anomalous Cosmic Rays or ACR in short.

The ACR oxygen ions are the most well studied for their intensity and spectral characteristics in the energy range 4–50 MeV/n. The intensity of ACR oxygen and nitrogen ions decreases from ~ 5 to ~ 50 MeV/n, where it merges with galactic cosmic rays (see Figure 1). The ACR helium ions have intensities higher than the GCR component in the 10–100 MeV/n energy range. Another striking feature of the ACR that was revealed, as the two spacecrafts, Pioneer 10 and 11, travelled further away from the Sun, was that the intensity of the ACR oxygen ions increased continuously and near the orbit of Jupiter it was ~ 40 times higher than that near the earth. These data established that the Sun cannot be the source of the anomalous cosmic rays and they must originate from sources or processes located beyond the orbit of Jupiter. So what could be the origin of this new type of cosmic rays? Several new ideas were proposed in the seventies to explain the presence of the ACR component. Before discussing these, we shall have a quick look at the results obtained from another experiment conducted in 1973–74 using the Skylab spacecraft.

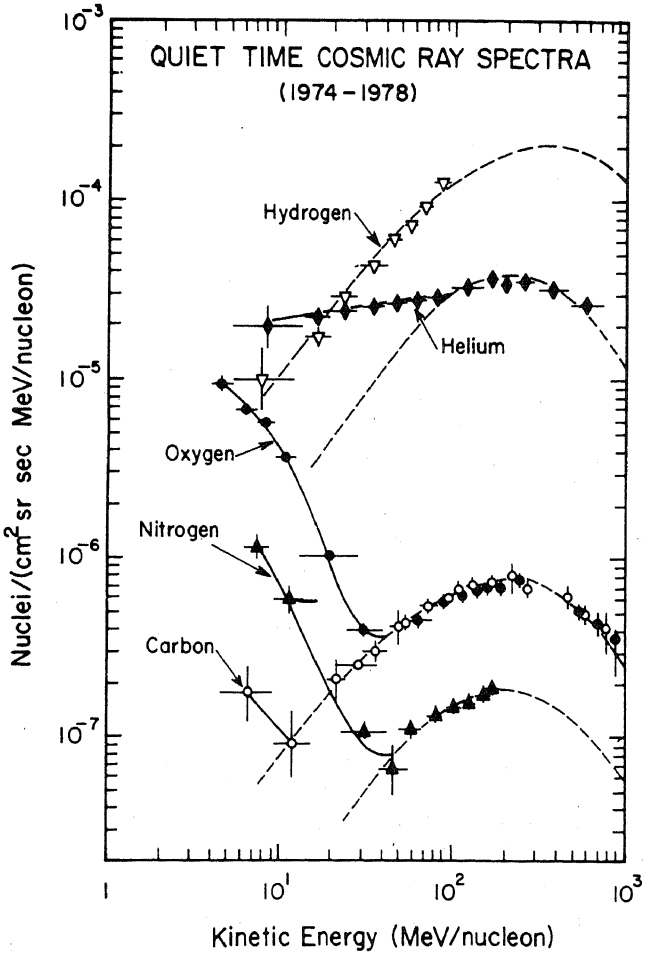


Figure 1. Energy spectra of H, He, C, N and O during solar quiet time of 1974–78. The solid lines indicate the ACR component; the dashed lines represent GCR component.

In the Skylab mission of NASA, a stack of passive solid state nuclear track detector (Lexan polycarbonate) was exposed on the outside of the Spacecraft for 74 days, between November 1973 and February 1974, a period that was entirely free from solar particle activity. Scientists from the Tata Institute of Fundamental Research (TIFR), Bombay, and the Physical Research Laboratory (PRL), Ahmedabad, have jointly analysed the data recorded in the detector stack and observed some strikingly new aspects of low-energy particles present within the earth's magnetosphere[5,6]. Firstly, ACR, O, N and C were observed for the first time inside the magnetosphere, with the same composition as in the interplanetary space. Secondly, it was concluded that these ACR particles must be in partially or singly ionized state; otherwise they will not be allowed by the earth's magnetic field to reach the Skylab orbit. Thirdly, the measured intensity of ACR oxygen was ~ 25 times higher than that expected on the basis of the measured interplanetary flux, from which it was concluded that ACR oxygen ions are trapped in the earth's magnetic field[6,7]. These findings made in 1975–80 have now been confirmed with instruments flown in different spacecrafts.

Where do anomalous cosmic rays come from? : Theoretical ideas

As the new data on the properties of the anomalous component of cosmic rays started accumulating, cosmic-ray theorists proposed a host of ideas to explain its origin. These ideas can be divided into two classes. The first category considers local galactic or solar system sources as responsible for generating the ACR component and predicts multiple charge states of ACR particles. McDonald *et al.*[2,8] suggested that the ACR particles originate in some unusual sources located in nearby interstellar space. Hoyle and Clayton[9] suggested that stars reaching white dwarf stage in stellar evolution may be responsible for the origin of ACR. Durgaprasad[10] proposed novae as possible sources, while Biswas *et al.*[11] suggested O-type stars emitting strong stellar wind as the ACR sources. A rather unusual model was proposed by Fowler *et al.*[12], who suggested that materials and gases brought in by comets are accelerated within the solar system to produce the ACR particles. All these theories would predict multiple charge states for the ACR particles, e.g. O^{1+}, O^{2+}, ... O^{8+}, etc. for ACR oxygen.

The other category of model, that considers local interstellar neutral atoms as the source of ACR, was proposed by Fisk *et al.*[13]. How do these neutral atoms evolve to become energetic anomalous cosmic-ray particles? A somewhat involved mechanism was proposed by the authors which constitute of four stages (Figure 2). In the first stage, interstellar neutral atoms enter the solar system as the Sun and the heliosphere (i.e. the volume in which solar plasma and magnetic field are confined) moves at a speed of 20 km/sec through the interstellar medium (ISM). The ionized particles in the ISM, on the other hand, are deflected at the heliospheric boundary by the solar magnetic field. This process therefore leads to the enhancement of those elements that exist predominantly in the neutral state of the ISM, i.e. helium, oxygen, nitrogen, and neon, and depletion of the elements that are mostly in ionized states, e.g. H, C, etc. In the second stage, when the interstellar neutral atoms approach the inner solar system, they are stripped off of their outermost electron by the strong solar UV radiation and they become *singly ionized* (i.e. He^{1+}, O^{1+}, N^{1+}, Ne^{1+}). This happens at different distances from the Sun, depending on the ionization potential of the element[14]. Once ionized, these particles are linked up with solar wind plasma and move outward with the speed of the solar wind, till they reach the boundary region of the heliosphere. The particles at this stage have an average energy typical of solar wind particles, i.e. 1 keV/n. In the third stage, these particles travel back and forth around the heliospheric shock front, which is the seat of random motion or turbulent magnetic field, and are accelerated from 1 keV/n to several hundred MeV/n. In the final stage, a fraction of the accelerated ACR ions diffuse back into the solar system where they suffer a decrease in their intensity and energy as they now move against the solar wind plasma in a situation similar to the case of transport of the more energetic galactic cosmic-ray particles. Thus we have singly ionized anomalous cosmic rays of ~ 1–100 MeV/n in the inner solar system. This model therefore predicts that ACR particles must be only in single ionized state, in contrast to the models of the first category which allows all ionization state of ACR. Thus, it was realized then that measurement of the ionization states of the anomalous cosmic-ray particles is crucial to pinpoint their source and origin. As this information was not available from experiments on board Pioneer 10/11 and IMP 8 spacecrafts there was a need for designing new experiments that could provide direct information on the ionization states of individual ACR particles. This brings us to the genesis of the Anuradha experiment.

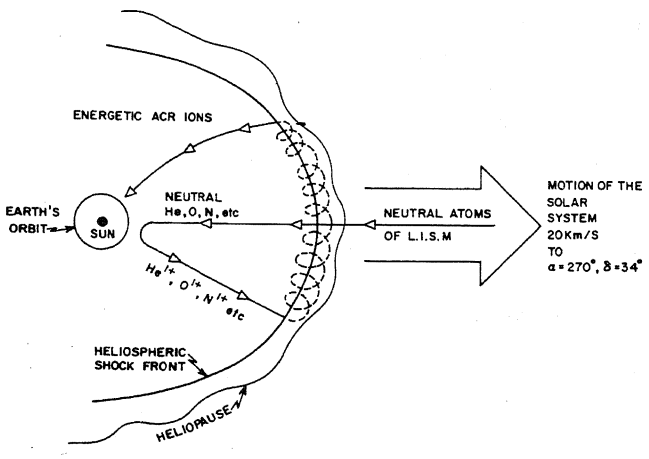


Figure 2. A schematic of the model proposed by Fisk *et al.*[13] to explain the origin of anomalous cosmic rays from interstellar neutral atoms.

Spacelab era begins: The 'Anuradha' cosmic-ray experiment

In 1977, the National Aeronautics and Space Administration (NASA) of USA invited proposals for new generation of scientific investigations from some of the major space research centres to be conducted in the Spacelab-1 mission aboard the Space Shuttle. The TIFR and PRL group of scientists, who were involved in the earlier Skylab experiment, jointly responded to the invitation and submitted a proposal entitled 'Studies of ionization states of solar and low energy galactic cosmic ray heavy nuclei', through the Indian Space Research Organization (ISRO). This proposal was selected by NASA for the Spacelab-1 mission, and this experiment, afterwards called 'Anuradha', was one of the two experiments from Asia, among about thirty US and European science experiments selected for this mission.

Scientific objective and instrument characteristics

The scientific objective of the Anuradha experiment was primarily to determine the ionization states of the newly discovered low-energy anomalous cosmic rays and also of solar energetic particles in case of solar energetic particle events occurring during the mission duration. The novel feature of this experiment was to use the earth's magnetic field as a giant spectrometer which filters out cosmic-ray particles according to their rigidity, i.e. momentum divided by effective charge (pc/Z^*). The instrument was designed in a manner that the arrival direction and time (and hence arrival location) of each of the cosmic-ray particle detected in the experiment could be determined. This information coupled with a model for the geomagnetic field allow us to trace back the probable complex trajectories of the incoming particles in the earth's magnetosphere and we can obtain

the minimum value of rigidity the particle must have to reach the Spacelab orbit at the point of its detection (see Figure 3). The other kinds of data obtained in the experiment are those recorded in the passive solid state nuclear track detectors which yielded the mass, nuclear charge, kinetic energy and momentum of each individual particles. Combining these results for each particle with its minimum permissible rigidity (R_{min}) obtained from the trajectory tracing calculation, the upper limit of ionic charge ($Z^* \leq pc/R_{min}$) of individual low energy anomalous and galactic cosmic-ray particles could be determined.

As the fluxes of the ACR particles are rather small, a large area detector of about 1200 cm^2 was used in this experiment. We used the newly discovered passive solid state nuclear track detectors with high sensitivity, called CR-39, and the detector films were assembled in the form of a circular stack composed of 150 sheets, each of thickness 250 μm and diameter about 40 cm.

A novel feature of the Anuradha experiment was its capability to determine the arrival time information of each cosmic-ray particles with an accuracy of ± 10 sec. This information and the record left by each particle in the detector, coupled with the information on the orientation and orbital parameters of the shuttle (which is available at all times) provide the arrival location (latitude, longitude and altitude) and direction of the particle as it entered the detector stack. To obtain the arrival time information, the detector module was divided into two stacks: a thin 'top' stack that was kept fixed and a thick rotatable 'bottom' stack that was coupled to a 15 bit absolute shaft encoder connected to a stepper-motor-gear assembly. The rotation of the bottom stack could be activated in steps of 40 sec of arc once in 10 sec so that it made a full rotation in 90 h. This part of the instrument was the most critical one, as very high mechanical precision in the fabrication and assembly is needed to achieve the high accuracy in the extremely small stepwise angular movement of a heavy load. The time information of an event in this detection arrangement can be obtained when an event is recorded both in the fixed 'top' and the rotating 'bottom' stacks of detectors. This is provided by the angular displacement between the two matched track segments produced by an individual event in the two detector stacks and the time history of stack rotation given by the shaft encoder output. Thus the information on the arrival time, location and direction of individual cosmic-ray particles is used for the trajectory computations and to obtain the value of R_{min} and finally its ionization states.

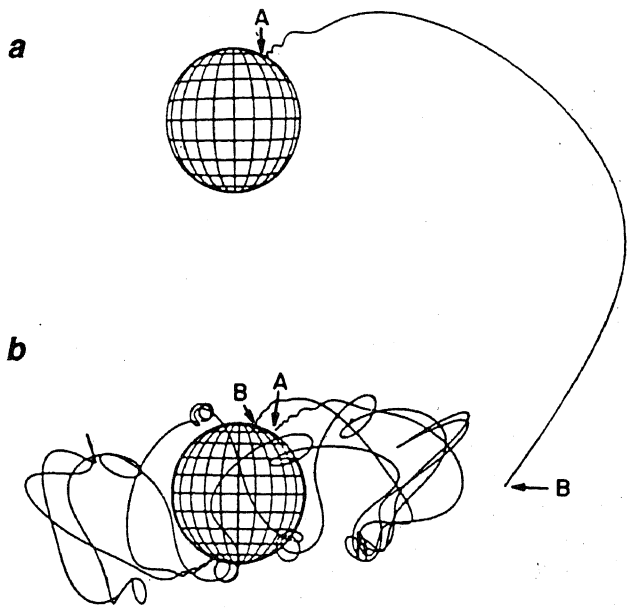


Figure 3. Complex trajectory of a cosmic-ray particle back-traced from its arrival point on Spacelab-3 orbit, for two values of rigidities: (*a*) shows an allowed trajectory that connects to the interplanetary space while (*b*) depicts the case when the assumed rigidity was too low for the particle to have an allowed trajectory.

One may ask the question why one needs the above complex procedures. The reason is that it is not possible to determine the ionization states of cosmic-ray particles by measurements confined in any detector system, active electronic or passive track detectors. This is so because when a cosmic-ray particle traverses a very small thickness of matter (~ a few μg/cm^2) in the upper most layer of a detector or its shielding, it is stripped off most of its orbital electrons and attain an equilibrium charge state and the information on the original ionization state is lost. Hence the momentum-filtering effect of the geomagnetic field is to be utilized, to obtain this information.

Anuradha takes shape

The detailed studies of experiment designs and requirements made in 1978–79 led NASA to work out a new schedule in 1980 for the Spacelab missions. It was arranged that Anuradha and another US experiment would be accommodated in Spacelab-3 mission in 1984–85, instead of in Spacelab-1. This rescheduling was most welcome as the orientation of Spacelab-3 was such that it provided continuous viewing of the deep space which was not the case of Spacelab-1.

A memorandum of understanding was signed by ISRO on behalf of the Govt of India and NASA in June 1980, according to which the Indian side had the responsibilities of the design, fabrication, assembling and testing of the instrument to meet the NASA specifications. NASA undertook its integration, flight operations, ground operations and related matters. The instrument and the data were to be returned to the Indian team for post-flight analysis.

The years 1980–82 were periods of intense activities on the design, procurement of materials and components, fabrication of mechanical, electrical and electronic systems and their integration, followed by various testings according to the NASA specifications. These were successfully accomplished by the engineering and scientific team members of TIFR–PRL–ISRO supported by BARC and several other centres. Six subgroups carried out specific works related to different subsystems. For example, the Detector Module Group of TIFR and the Laser Group of BARC had set up a special arrangement of CO_2 laser and He–Ne laser to cut ~ 500 CR-39 detector sheets—which are glass-like and brittle—into annular shapes of required dimensions. About 150 sheets of these were assembled in stainless steel housing and three such units were constructed.

The top enclosure—a curved sheet of Al alloy of thickness ~ 100 μm—was a critical element of the instrument as it must be thin enough to allow low energy (~ 8 MeV/n oxygen nuclei) to enter the detector and at the same time thick enough to provide an air-tight enclosure. Two pairs of venting valves maintained the inside pressure of the instrument at 0.1 atm during the flight and at 1 atm at all other times. The electrical and electronic systems and necessary softwares were fabricated and tested by the PRL group. This subsystem provided the required pulses to the high resolution stepper motor, which rotated the bottom stack by an increment of 40 sec of arc once in 10 sec (or by any other predetermined rate) and has the capability to correct for any overshoot or undershoot in the step movements. The rotational motion was read out by the 15 bit absolute encoder which was displayed and was monitored along with house keeping data such as the inside temperature of the instrument. Details of the instrument design and operation are given elsewhere[15] and a summary of the main characteristics is given in Table 1. A sectional view of the Anuradha instrument is shown in Figure 4.

Various tests of the instrument were carried out according to the NASA requirements at different Indian centres. The random and sinusoidal vibration tests, the thermovacuum and electromagnetic tests were carried out at the ISRO Satellite Centre, Bangalore. The acceleration tests were performed at the ISRO Shriharikota Centre and acoustic tests were done at IIT, Madras.

Four units of the Anuradha instrument were constructed—the engineering model for design testing, prototype model for all tests qualifications, the flight instrument for actual flight and the flight standby for contingency. After NASA's acceptance of test records, the flight instrument and the flight standby—installed

Table 1. Summary of Anuradha instrument features and flight characteristics

Weight	:	50 kg
Volume	:	48 cm diameter; 56 cm height
Detector	:	Composite CR-39 detector modules coupled to a high resolution stepper motor and a 15-bit absolute optical shaft encoder for arrival time and direction measurements of cosmic ray ions
Spacelab flight	:	Space shuttle Challenger launch – 29 April 1985; 1602 h GMT, and landing 6 May 1985; 1611 h GMT
Orbit	:	Altitude : 352 km Inclination : 57° to the equator Attitude : Gravity gradient stabilization
Ions exposure	:	Instrument activation, 123-12-44 GMT and deactivation, 126-05-00 GMT
Power allocation	:	10 W average for 90 h
Thermal control	:	Cold plate in pallet and a multilayer insulation
Command and data management	:	5 discrete on-off command channels 16 discrete data acquisition channels 1 analogue data channel
Experiment control	:	By stored program in on-board computer of SL-3 through I/O unit and interfacing with remote acquisition unit (RAU)
Data acquisition	:	CDMS telemetry of 372 bits every 1 s via TDRS and other satellites to ground station for stack movement and other house keeping data. Other data are recorded in track detectors and are processed post-flight.

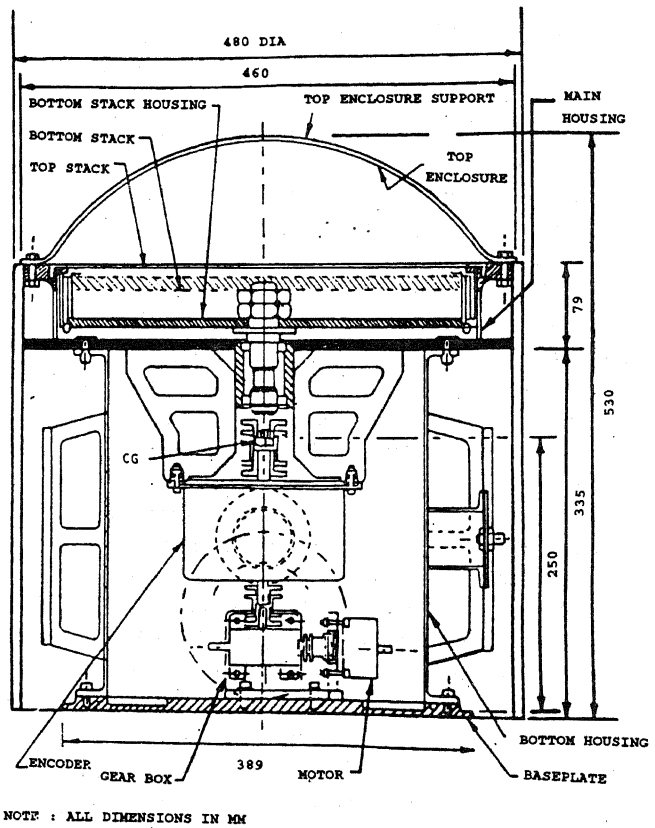


Figure 4. A sectional view of the 'Anuradha' cosmic-ray instrument.

specially designed containers and other accessories—were airlifted from India to NASA Kennedy Space Center, Florida, USA in December 1983. During 1984, NASA and Indian engineers installed Anuradha in the Spacelab-3 on the instrument-support structure and integrated it with Spacelab-3 electrical, electronic software and thermal control systems. In addition to passive thermal control with multilayer insulation, Anuradha was provided with active thermal control through a cold plate with circulating freon, on which the instrument was mounted. After the completion of various tests of the integrated payload, Spacelab-3 aboard Space Shuttle Challenger was ready for spaceflight in December 1984.

Space shuttle in orbit: Anuradha collects cosmic data

On mid-day of 29 April 1985, the space shuttle Challenger with Spacelab-3 on board was launched successfully into a circular orbit at a height of 350 km at an orbital inclination of 50° to the equator. The thirteen experiments flown in Spacelab-3 were in the areas of material science, life science, fluid mechanics, atmospheric science and astrophysics; countrywise, ten of these belonged to USA, two to France and one to India. A special feature of the SL-3 mission was the role of crew members, seven in number, five of whom were specially trained for the science experiments. A vax computer provided on-board support for control and operation of the experiments. These and other facilities in the mission enabled the scientists to undertake several new types of experiments for the first time. The Anuradha experiment could be activated and controlled either by SL-3 computer by a crew member or by ground commands from the control centre at Houston, Johnson Space Centre (JSC). Data were transmitted from Spacelab-3 by relay satellites to JSC Houston and Goodard Space Flight Center, Maryland.

When the Anuradha experiment was activated on the third day of the mission according to schedule it was found that some command channels were not responding while some others were operating. After studying the instrument response the Indian team informed the mission manager that the instrument characteristics are alright and the problem must be somewhere else, which should be identified and rectified. After a series of tests conducted by the NASA ground engineers and SL-3 crew, the problem was identified in the faulty NASA cable connecting SL-3 computer to the Anuradha experiment. As there was no spare cable, this problem was ingeniously solved by NASA engineers by replacing the faulty cable by a cable of another experiment, which just completed its operation. These and associated activities were carried out during a 24 h period and the instrument was switched on again on 3 May 1985. It was an exciting and dramatic moment for all present at the control centre to see that all channels of Anuradha were responding beautifully and the instrument operated perfectly for the next three days. It was switched off at the experiment closing time of the mission on 5 May 1985. A few photographs of the Anuradha and SL-3 mission are shown in Figure 5.

A schematic diagram of the experiment operation in orbit and of the flow of signals between Anuradha instrument in Spacelab-3 and the ground control centre is shown in Figure 6. The instrument operation takes place in the following manner. Once every 10 sec command signal generated by the experiment computer rotates the 'bottom' detector stack by one step of 40 arcsec and the encoder output is read out. The encoder reading, together with GMT, instrument temperature and other relevant information are transmitted from the SL-3 by high-rate multiplexer to payload operation and control centre once in 2 sec. The stepwise rotation is repeated once in every 10 sec, until a maximum of 32,736 rotational steps are made during about four-day period. In actual operation, Anuradha experiment had undergone rotational steps for about 64 h, i.e. for ~ 70% of the maximum value. The telemetered data of the instrument and of the Spacecraft position and orientation, were both recorded every 2 sec in the data tapes at the control centre and these tapes containing about 400 million data bits were provided to us by NASA for postflight analysis.

Postflight activities, data acquisition and synthesis

The postflight operations included the calibration of CR-39 stacks with three types of accelerator beams. Firstly, 50 MeV alpha particle beams from the Variable Energy Cyclotron (VEC), Calcutta were used to provide visible markings of the starting position of rotation on both the stacks. Secondly, 1 GeV/n ^{179}Au beams from the Bevalac, University of California at Berkeley, were used in vertically incident manner at six positions of the bottom stack to facilitate realignment of the CR-39 sheets after disassembly and processing. Thirdly, 140 MeV/n ^{56}Fe beams incident at 45° were used in six positions in the bottom stack to provide known events for charge calibration of the CR-39 detector.

Energetic cosmic ray nuclei incident on solid state track detectors such as CR-39 cause radiation damage along their path by breaking the chemical bonds of the constituent molecules. This leads to the formation of nascent damage trails of diameter ~ 100 Å which are preserved as latent tracks. When CR-39 detector sheets are treated with strong alkaline solution (e.g. 6.25 N NaOH at 70°C for 6 h) preferential etching takes place along the latent damage trails, which are revealed as conical etched tracks visible in an optical microscope.

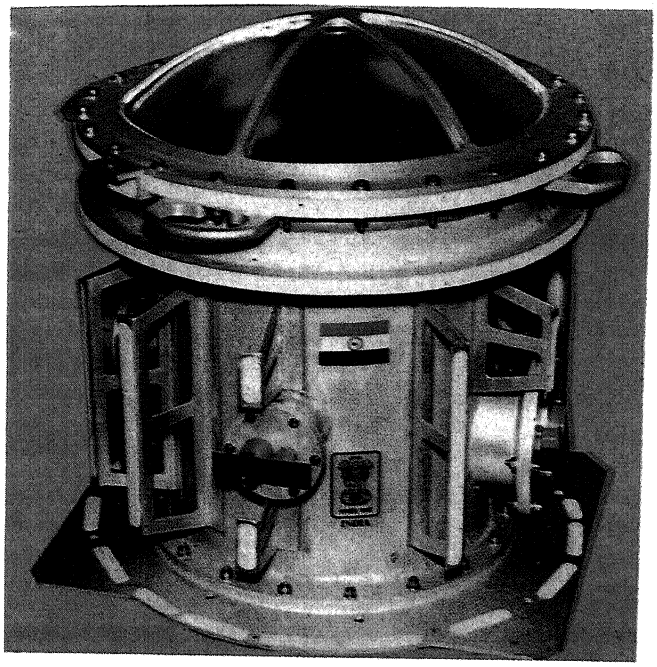


Figure 5 *a*. The Anuradha instrument.

The length of the etched cones, which is a measure of the ionization loss rate, and the residual range can yield the nuclear charge and kinetic energy of the incident particle[16,17]. Etched cones of a cosmic ray Fe nucleus traversing several CR-39 sheets are shown in Figure 7. In Anuradha experiment we initially scanned an annular area of width 2 cm in the two uppermost CR-39 sheets using a specially constructed semi-automatic microscope whose stage could be rotated in angular steps of 40 arcsec and in radial direction, R. The selected tracks of cosmic-ray particles, located in R, θ coordinate system, were followed through successive plates in the stack till the end of their range. A detailed analysis of track parameters was made which yields the atomic number, kinetic energy and other particulars of the incident cosmic-ray particles. Thus cosmic-ray ions of nitrogen to nickel of 8 to 150 MeV/n were identified.

The procedure of determining the arrival time and direction of the particles involves more complex procedures as briefly described earlier. By matching track segments in the top (fixed) stack and in the bottom (rotating) stack using algorithm that uses four or more track parameters, angular displacement between the track segments is obtained. This information when combined

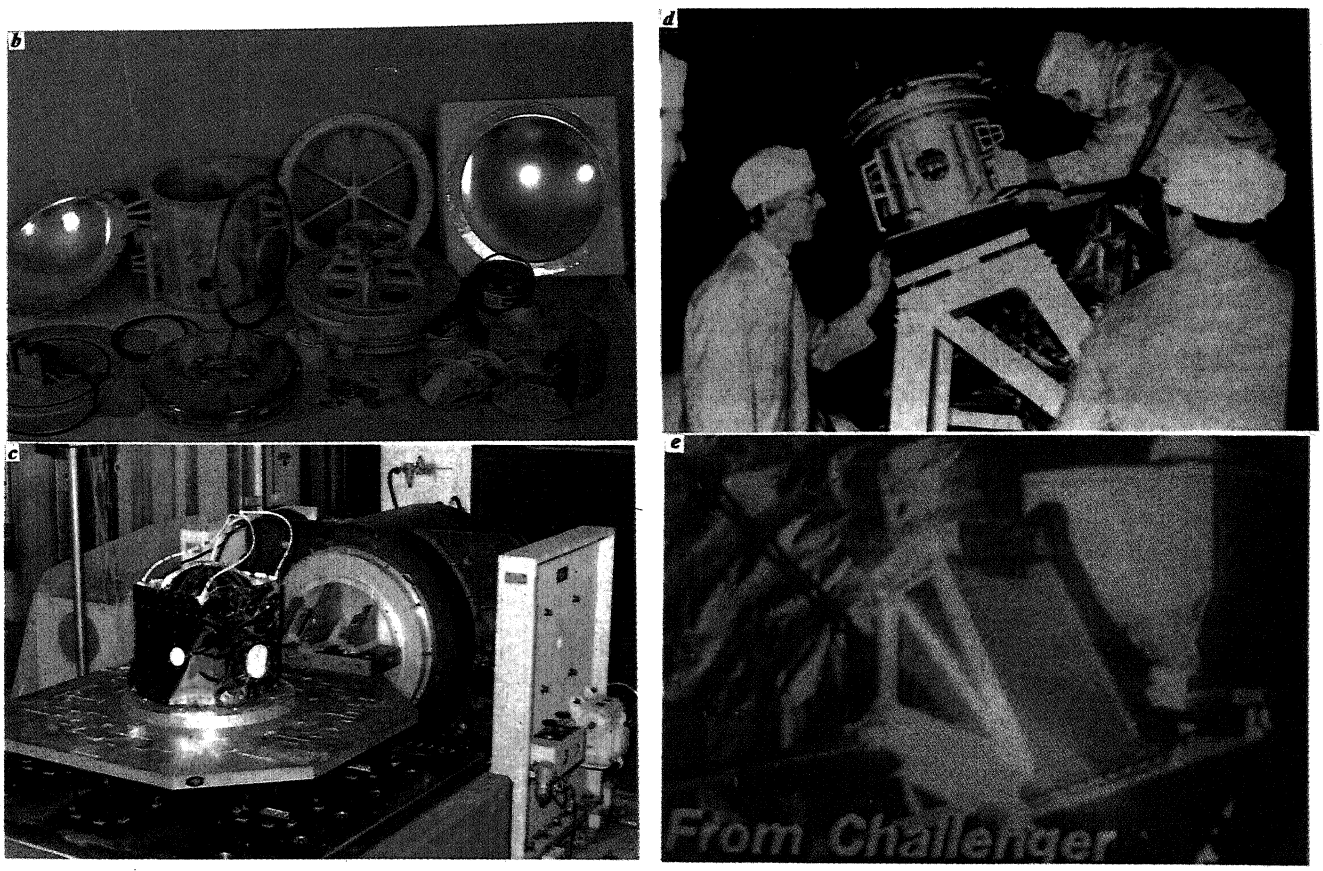


Figure 5 *b–e*. ***b***, The Anuradha instrument: individual components; ***c***, The Anuradha instrument on the vibration testing platform; ***d***, The Integration of Anuradha to Spacelab-3 platform; ***e***, Anuradha in space (on board space shuttle Challenger).

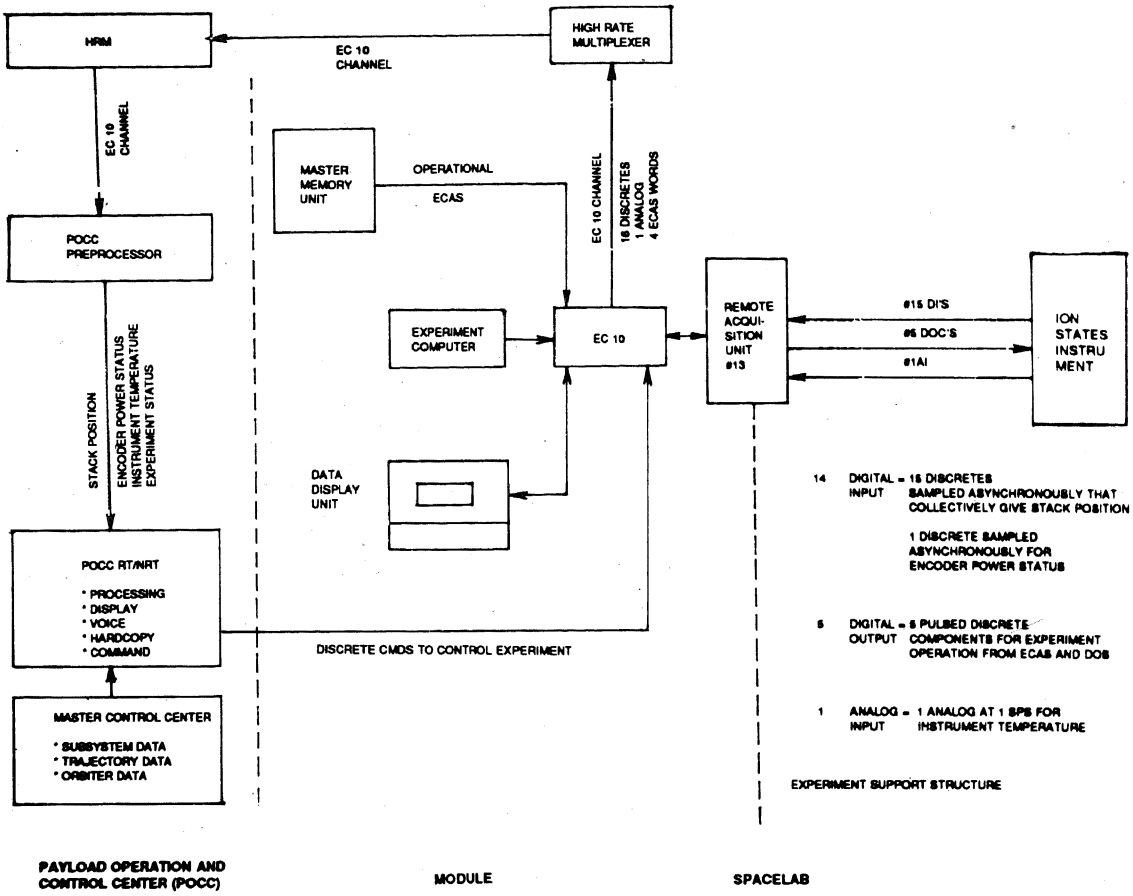


Figure 6. Block diagram showing the experiment operation and flow of signals between the Spacelab module, the instrument and the ground control centre.

with the telemetered data on encoder position vs GMT and orbital parameters of SL-3 yielded the arrival time and direction of the cosmic-ray particle.

Complete details of these procedures are given by Singh[18] and Dutta[19]. A total of about 100 cosmic-ray particles of carbon to nickel in the energy range ~ 12 to 150 MeV/n were identified and all the relevant data obtained in each case. For the trajectory tracing calculations we have followed the approach of Shea and Smart[20] and used a polynomial expansion of 10th order and the parameters suggested by the International Geomagnetic Reference Field (IGRF) for the epoch 1985 for an accurate representation of the geomagnetic field. The effects of external current system etc. were also considered in the final analysis of the data.

Anuradha results: New findings in anomalous and galactic cosmic rays

The scientific results from the Anuradha experiment had several new features and we shall briefly discuss these and their implications.

Firstly, the Anuradha experiment established the detailed method of using the earth's magnetic field as a giant momentum analyser for the accurate determination of minimum rigidities of individual cosmic-ray particles utilizing the arrival time and direction information[21,22].

Secondly, the mean ionization state of a sample of ACR oxygen ions was determined as O^{+1} for the first time using the orbit averaged flux method[23,24]. In this method data for the interplanetary ACR flux obtained by spacecrafts are used to calculate the expected flux of ACR oxygen inside the magnetosphere for different assumed ionization states. These values were then compared with the measured orbit averaged fluxes in Anuradha to determine the charge state of ACR oxygen as 1^+ (Figure 8). Recent studies by a Russian–US group of scientists, who used this method and analysed data from COSMOS and IMP-8 satellites, have confirmed this result[25].

Thirdly, the ionization states of 75 individual low-energy GCR and ACR particles were successfully measured for the first time[21,26]. Sixteen of these events were identified as ACR particles of N, O and Ne, whose upper limit of ionization states show that they are singly ionized or consistent with it[21,26] (Figure 9). The distributions in geographic latitude and longitudes of these ACR particles were found to be uniform (Figure 10) as expected and so also their arrival times.

Fourthly, a new unexpected result was obtained while measuring the ionization states of low-energy heavy ions. It was found that galactic cosmic-ray ions of sub-iron (Sc to Cr) and iron (Fe–Ne) group of 30–150 MeV/n are composed of about 20% particles which are partially ionized, e.g. Ti^{6+}, Cr^{5+}, Fe^{10+}, etc.[22,26]. This is contrary to the normal assumption that all GCR particles are fully ionized.

Lastly, another significant observation is that the abundance ratio of sub-iron to iron group of particles in low-energy GCR (30-130 MeV/n) within the magnetosphere is about a factor of two higher than the interplanetary value[27]. This, we believe is due to the presence of partially ionized ions in GCR at these energies.

Implications of the new findings, gateways to the galaxy and future perspective

The results from the Anuradha experiment established that ACR particles originate from interstellar neutral atoms and these are in singly ionized state as proposed by Fisk *et al.*[13]. However, details of the acceleration mechanism are yet to be fully understood[28]. It is gratifying to note that recent satellite experiments of German–US collaboration[29] and Caltech–NASA groups[30] have adopted the basic method of the Anuradha experiment and have confirmed our results on the ionization states of ACR particles.

The ACR particles bring to us a new sample of interstellar matter undergoing ionization and acceleration. One can attempt to deconvolute the processes to estimate the abundances of neutral atoms in ISM as discussed elsewhere[28]. Also, one can determine the isotopic composition of neutral atoms in ISM from ACR particles, as these remain unaffected by the intervening processes. A recent satellite experiment has also identified the presence of a third radiation belt composed of trapped ACR particles in near earth space[30]. This confirmed our observations made in the Skylab experiment in 1973–74 (see review by Biswas[31]). Now we can have easy access to a large flux of interstellar particles in the near-earth space.

Energetic particles of galactic cosmic rays (GCR) have travelled a significant amount of matter (~ 10 g cm^{-2}) in interstellar medium before reaching us. Hence these are mostly bare nuclei, stripped of their orbital electrons. The known values of electron capture and

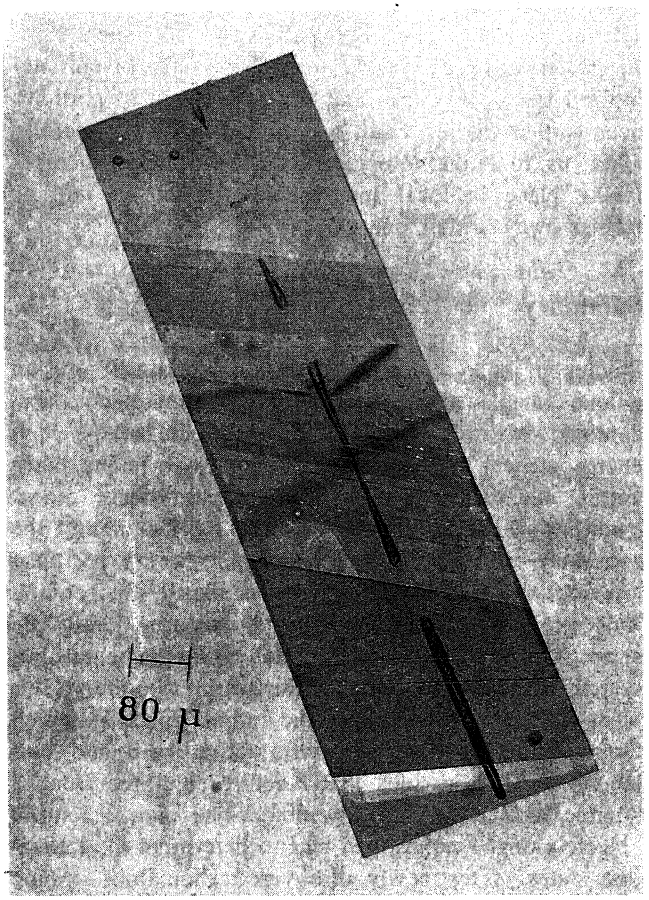


Figure 7. Nuclear track produced by a Fe ion in several sheet of CR-39 detector used in the Anuradha experiment.

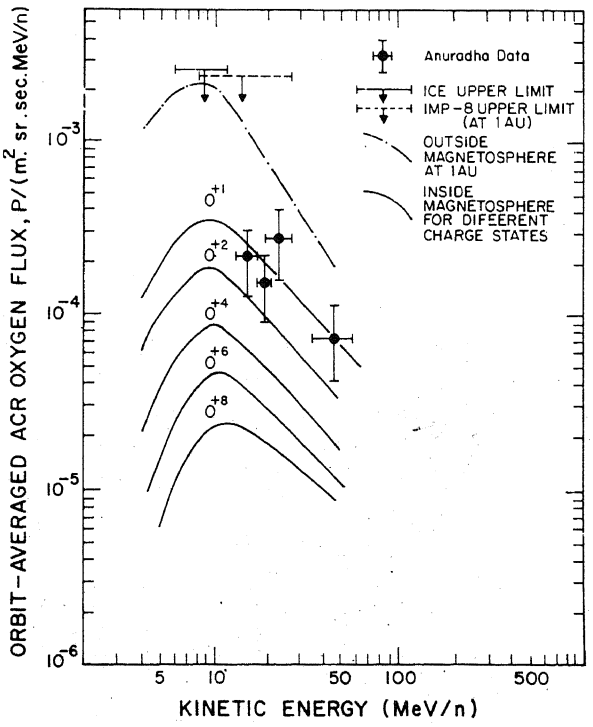


Figure 8. Measured fluxes of ACR oxygen in Anuradha experiment and the expected values, based on interplanetary fluxes from ICE, IMP and Voyager data for different charge states. The Anuradha data suggest ACR oxygen to be singly ionized.

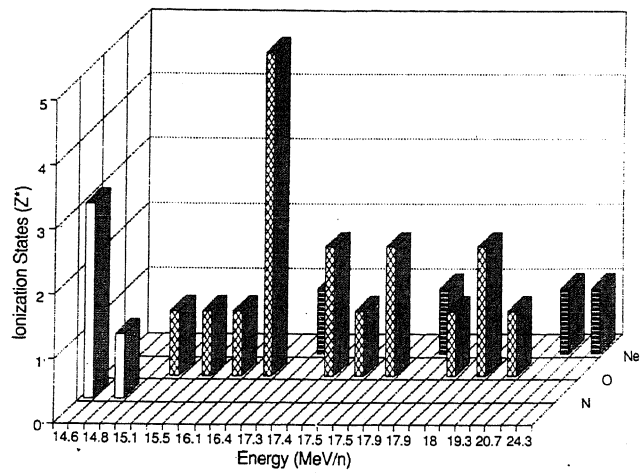


Figure 9. The measured upper limits of ionization states of ACR particles of N, O and Ne as a function of kinetic energy.

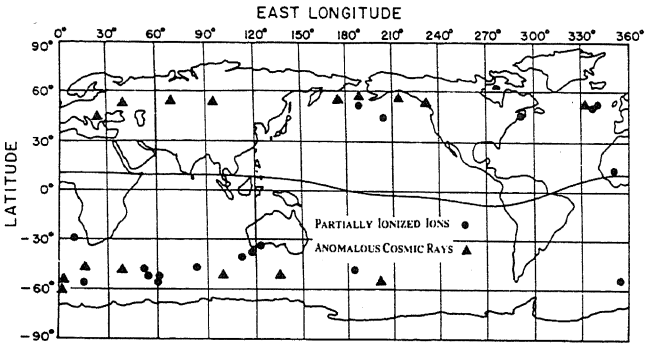


Figure 10. The geographical distributions in the latitudes and longitudes of the arrival locations of the ACR and partially ionized GCR particles.

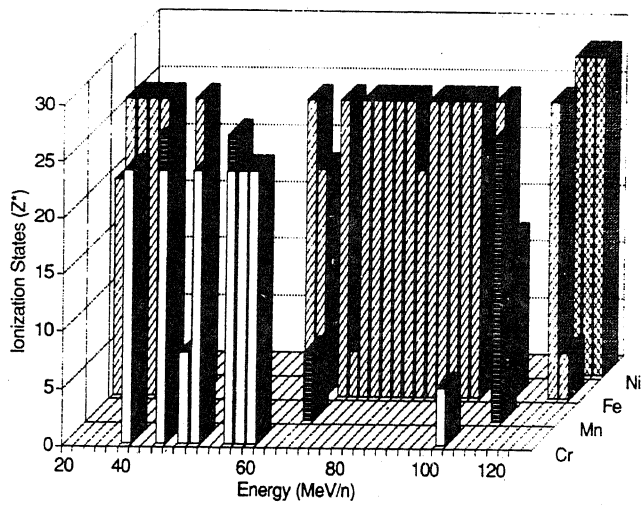


Figure 11. The measured upper limits of ionization states of low energy GCR heavy ions (Cr-Ne) detected in the Anuradha experiment. Note the presence of partially ionized particles.

stripping cross-sections measured in the laboratory show that iron group nuclei in GCR can capture orbital electrons only at very low energies (1–5 MeV/n) and above 10 MeV/n, electron stripping dominates. Therefore, the important question arises, how do we get GCR iron and sub-iron ions of ~ 100 MeV/n in partially ionized states observed in Anuradha? Logical conclusion appears to be that these particles captured electrons at low energies and are subsequently accelerated to > 100 MeV/n in ISM and have traversed very small amount of matter before reaching the solar system. Thus, our observation provides evidence for reacceleration or distributed acceleration of GCR in ISM. Probable sites for reacceleration are shock fronts of supernova remnants in ISM. Confirmation of the above will have important new implications on acceleration of galactic cosmic rays. Several attempts are currently on to study this enigmatic component of cosmic rays (e.g. see ref. 32).

Future space experiments may explore these new pathways to interstellar medium using low energy anomalous and galactic cosmic rays as tracers. Measurements of ACR oxygen flux beyond the orbit of Uranus during the period of minimum solar activity in 1997–98 will give definitive results on the intensities of these particles at the boundary region of the heliosphere and hence their energy density in the local ISM. In galactic cosmic rays, the new phenomena of partially ionized low energy iron and sub-iron group of particles need further studies. Such experiments should hopefully allow us to build new gateways to explore processes taking place in local interstellar medium as well as further away within our galaxy.

Summary

Studies of energetic cosmic-ray particles, that are an unique sample of matter from the interstellar space, continue to provide us with new insights of the high energy nuclear and astrophysical processes taking place within our galaxy and beyond. A significant recent contribution in this context has been the findings of the extremely successful Indian cosmic-ray experiment 'Anuradha' flown on board space shuttle Spacelab-3. This experiment was planned and executed jointly by scientists from the Tata Institute of Fundamental Research, Bombay and Physical Research Laboratory, Ahmedabad, under the aegis of the Indian Space Research Organization (ISRO), and with active help and collaboration from various ISRO centres and the Bhabha Atomic Research Centre, Bombay. The main scientific objective of the experiment was to determine the ionization states of the newly discovered anomalous cosmic rays to identify their source and origin. A very special experimental payload, that combines a passive nuclear particle detector and an active electro-mechanical system,

was designed and constructed to achieve this objective. Some of the important results obtained in this experiment are: (i) a clear demonstration that the geomagnetic field can be used as a giant momentum analyser for determining the rigidity (momentum/charge) of individual low energy cosmic-ray particles impinging on the earth, (ii) the first direct determination of the ionization states of individual anomalous cosmic-ray particles, which showed them to be singly ionized and confirmed the interstellar neutral atoms as their source, and (iii) detection of low energy partially ionized iron group particles in galactic cosmic rays, a completely unexpected result with important significance. The basic principle used in the Anuradha experiment has been adopted in a couple of recent satellite experiments that have confirmed the results on the ionization states of the anomalous cosmic rays obtained in the Anuradha experiment. Further studies of the enigmatic partially ionized iron group particles are also being pursued in these experiments. The findings of the Anuradha experiment point towards the exciting possibility of understanding the elemental and isotopic composition of the local interstellar matter by sampling the anomalous cosmic rays in the interplanetary or the near-earth space.

1. Wolfendale, A. W., in Proceedings of XXIII International Cosmic Ray Conference (eds Sonnet, C. P. *et al.*), World Scientific, Singapore, 1993, pp. 143–154.
2. McDonald, F. B., Teegarden, B. J., Trainor, J. H. and Webber, W. R., *Astrophys. J. (Lett.)*, 1974, **187**, L105–108.
3. Hovestadt, D., Vollmer, O., Gloeckler, G. and Fan, C. Y., *Phys. Rev. Lett.*, 1973, **31**, 650–653.
4. Garcia-Muroz, M., Mason, G. M. and Simpson, J. A., *Astrophys. J. (Lett.)*, 1973, **182**, L81–84.
5. Biswas, S., Durgaprasad, N., Nevatia, J., Venkatacardan, V. S., Goswami, J. N., Jayanthi, U. B., Lal, D. and Matto, S. K., *Astrophys. Space Sci.*, 1975, **35**, 337–347.
6. Biswas, S. and Durgaprasad, N., *Space. Sci. Rev.*, 1980, **25**, 285–327.
7. Blake, J. B. and Freisen, L. M., Proceedings of the 15th International Cosmic Ray Conference, 1977, vol. 2, pp. 341–344.
8. McDonald, F. B., Lal, N., Trainor, J. H., Van Hollebeke, M. A. I. and Webber, W. R., *Astrophys. J.*, 1977, **216**, 930–939.
9. Hoyle, F. and Clayton, D. D., *Astrophys. J.*, 1974, **191**, 705–710.
10. Durgaprasad, N., *Astrophys. Space Sci.*, 1977, **47**, 435–445.
11. Biswas, S., Durgaprasad, N. and Trivedi, S., *Proc. Indian Acad. Sci. (Earth Planet. Sci.)*, 1981, **90**, 337–344.
12. Fowler, P. H., Redfern, R. M. and Swordy, S. P., *Nature*, 1979, **279**, 622–624.
13. Fisk, L. A., Kozolovosky, B. and Ramaty, R., *Astrophys. J. (Lett.)*, 1974, **190**, L35–37.
14. Axford, W. I., in *Solar Wind-III* (eds Sonnet, C. P., *et al.*), pp. 609–660.
15. Biswas, S. *et al.*, *Pramana*, 1986, **27**, 89–104.
16. Fleisher, R. L., Price, P. B. and Walker, R. M., *Nuclear Tracks in Solids*, Univ. of California Press, Berkeley.
17. Henke, R. P. and Benton, E. V., *Nucl. Instrum. Methods*, 1971, **97**, 483–489.
18. Singh, R. K., Ph D Thesis, Bombay Univ., Bombay, 1990.
19. Dutta, A., Ph D thesis, Gujarat Univ., 1992.
20. Shea, M. A. and Smart, D. F., *AFGRL-TR-75-0185*, 1975 (also private communication, 1988).
21. Singh, R. K., Mitra, B., Durgaprasad, N., Biswas, S., Vahia, M. N., Yadav, J. S., Dutta, A. and Goswami, J. N., *Astrophys. J.*, 1991, **374**, 753–765.
22. Biswas, S., Durgaprasad, N., Mitra, B., Singh, R. K., Dutta, A. and Goswami, J. N., *Astrophys. J. (Lett.)*, 1990, **359**, L5–9.
23. Biswas, S., Durgaprasad, N., Mitra, B., Singh, R. K., Vahia, M. N., Yadav, J. S., Dutta, A., Goswami, J. N., *Astrophys. Space Sci.*, 1988, **149**, 357–367.
24. Dutta, A., Goswami, J. N., Biswas, S., Durgaprasad, M., Mitra, B. and Singh, R. K., Proceedings of the 22nd International Cosmic Ray Conference, 1991, vol. 3, pp. 346–349.
25. Adams, J. H., Jr. *et al.*, Proceedings of the 22nd International Cosmic Ray Conference, 1991, vol. 3, pp. 358–361.
26. Dutta, A., Goswami, J. N., Biswas, S., Durgaprasad, N., Mitra, B. and Singh, R. K., *Astrophys. J.*, 1993, **411**, 418–430.
27. Biswas, S., Durgaprasad, N., Singh, R. K., Vahia, M. N., Yadav, J. S., Dutta, A. and Goswami, J. N., *J. Astron. Astrophys.*, 1994, **15**, 85–94.
28. Biswas, S., Durgaprasad, N., Mitra, B. and Dutta, A., *Space Sci. Rev.*, 1993, **62**, 3–65.
29. Klecker, B., McNab, M. C., Blake, J. B., Hamilton, D. C., Hovestadt, D., Kastle, H., Looper, M. D., Mason, G. M., Mazur, J. E. and Scholer, M., *Astrophys. J. (Lett.)*, 1995, **442**, L69–72.
30. Cummings, J. R., Cummings, A. C., Mewaldt, R. A., Selesnick, R. S., Stone, E. C. and Von Rosenvinge, T. T., Proceedings of the 23rd International Cosmic Ray Conference, 1995, vol. 3, pp. 428–431.
31. Biswas, S., *Space Sci. Rev.*, 1995 (in press).
32. Tylka, A. J., Boberg, P. R., Adam, J. H. Jr., Kleis, T. and Beaujean, R., *Astrophys. J. (Lett.)*, 1995, **438**, L83–86.

Cosmic gamma ray bursts – Recent developments and observations from SROSS satellites

K. Kasturirangan, V. N. Padmini, N. L. Prasad and U. R. Rao

ISRO Satellite Centre, Airport Road, Vimanapura P.O., Bangalore 560 017, India

Amongst a dozen energetic phenomena discovered during the past quarter century, the enigma of classical gamma ray bursts (GRBs) is as great as ever, with no consensus on what causes these bursts. The once dominant galactic neutron star model of GRBs has lost its popularity to the extragalactic and cosmological models in view of COMPTON's observations. The observational efforts in tracking down these mysterious sources are at the moment truly international, with emphasis on the best localizations followed by deep searches for counterparts at other wavelengths.

COSMIC gamma ray bursts (GRBs), first reported by Klebesadel *et al.*[1], continue to remain as one of the long-standing puzzles in high energy astrophysics, after twenty-five years of their discovery. GRBs are transient, non-terrestrial phenomena occurring randomly in space and time, lasting a few ms to a few minutes and characterized by bursts of intense radiation, predominantly at gamma ray wavelengths with bulk of their emission > 100 keV. The US Vela satellites that were launched to monitor any violations of the 1963 nuclear test ban treaty, recorded these events on their gamma ray detectors as early as 1967, but the announcement of this discovery of a new class of astrophysical phenomena had to wait until 1973, when the US military allowed the information to be made public.

During the period from 1973 to 1991, over 500 bursts with fluence $>10^{-6}$ ergs/cm^2 have been recorded by instruments mostly scintillation spectrometers flown onboard at least 20 satellites, viz. Vela, Helios 2, HEAO 1 & 2, Prognoz, ISEE-3, Solar max mission, Ginga, Apollo, etc. and inter-planetary probes like Venera 11, 12, 13, 14, Pioneer Venus orbiter, Phobos, etc. Studies on GRBs have proved difficult due to their transitory nature. The most general methodology of GRB study involves temporal and spectral studies, spatial distribution of an ensemble of bursts and other wavelength observations. Data collected on GRBs usually consist of a record of intensity variations as a function of time, over a wide energy band (Temporal profile) and energy spectra recorded at constant time intervals.

Over the years, two classes of GRBs have been recognized[2], viz. Classical GRBs, which account for the majority of bursts observed and Soft Gamma Repeaters (SGRs), three of which are known. The relation between the two, if any, is not clear. Time profiles of classical GRBs display a wide variety of shapes and durations varying from few seconds to a hundred seconds, with hard spectra (peak power at energies > 100 keV) which evolve with time. Typical rise and decay times range from 0.1 to 1 s. Classical GRBs are not known to repeat. SGRs, have nominal durations of 0.1–1 s, exhibit soft spectra with e-folding energy ~ 30 keV, which are time invariant. SGRs repeat at irregular intervals.

Although about a hundred models have been proposed to explain the origin of gamma ray bursts, none has so far been able to explain all the observations satisfactorily. The mystery shrouding soft gamma repeaters has been of late, largely resolved by multiwavelength counterpart observations. In contrast, classical GRBs have so far not been identified with any known population of objects in the universe, despite conducting deep searches to identify their counterparts at other wavelengths. As the distance to the source is as yet unknown, the total energy involved in the burst process has remained ambiguous by about 23 orders of magnitude.

With the launch of NASA's Compton Gamma Ray Observatory (CGRO) in April 1991, which carries onboard an instrument to monitor GRBs with unprecedented sensitivity, viz. the Burst and Transient Source Experiment (BATSE)[3], there has been a significant change in the quality of data and also the interpretation of the results. BATSE with its sensitivity to bursts with fluence $>5\times10^{-8}$ ergs/cm^2, which is an order of magnitude superior to its predecessors, records about one burst per day and has accumulated data on about 1296 bursts till 8 June 1995 (ref. 4). Prior to Compton, there existed a wide consensus that gamma ray bursts originate due to some catastrophic event on or near the surface of a neutron star in our galaxy. The discovery by BATSE that bursters have an isotropic, but radially inhomogenous distribution in space has altered the thinking of many theorists. As CGRO observations favour a distribution greater than at least 50 kpc, the mystery on the origin of GRBs has only deepened further and at the moment it is unclear whether these are of galactic or extragalactic origin. Recent reviews on GRBs, by Hurley[5], Higdon

and Lingenfelter[6] and conference proceedings edited by Fishman *et al.*[7], provide abundant information on observations, analysis and models of GRBs.

In the following sections, we review GRBs, from an observational point of view, as they were understood prior to the launch of CGRO and the highlights of results following the launch. The motivation for the Indian GRB experiments flown on the Stretched Rohini Satellite Series (SROSS-C, May 1992; SROSS-C2, May 1994) and their observations are also presented. Also discussed is the possible contribution of SROSS-C2 to the third interplanetary network of detectors currently monitoring GRBs.

Pre-CGRO era

Prior to the launch of CGRO, there existed a wide consensus that a GRB is caused by some catastrophic event near the surface of a neutron star in our galaxy. The observations suggested that most of the common classical burst sources are single, intensely magnetic neutron stars in the nearby galactic disk, and that the SGRs could be neutron stars with slightly weaker magnetic fields and possible binary companions. The circumstantial evidence which led to the galactic neutron star hypothesis was:

i) Very small source size < 1000 km, implied by rapid time variability in classical GRBs and < 60 km implied by the 0.2 ms rise time of GB790305b (ref. 8).
ii) Location of GB790305b event within the N49 supernova remnant (SNR) in the Large Magellanic Cloud (LMC) (ref. 8).
iii) Very high mean density > 10^6 g/cc implied as minimum density against breakup, if the 8 s period in the soft repeater GB790305b is rotational[9–11].
iv) Absorption lines between 20 and 70 keV in about 15% of classical bursts, which are attributed to cyclotron absorption in Terragauss magnetic fields[12,13]. These were further confirmed by the Japanese satellite Ginga[14] which observed bursts with two lines harmonically spaced, viz. at 20 & 40 keV.
v) Emission lines seen between 350 and 500 keV in both classes of bursts[12,13] – interpreted as gravitationally redshifted electron position annihilation radiation with neutron star redshifts of 0.1 or more[15].
vi) Failure to detect quiescent counterparts at other wavelengths within the source positional error boxes[16–22].
vii) Galactic distances implied by the black body limit derived from soft X-ray emission from GB870303 (ref. 14).

Further, the size-frequency distribution, which indirectly provides information on the radial distribution of sources, was found to follow the $-3/2$ power law for bright bursts, while the distribution deviated significantly for weak bursts with a flattening at lower fluences. This flattening at low fluence was considered as an evidence of a galactic origin, with more distant sources exhibiting nonuniform spatial distribution. The angular distribution of sources was however found to be isotropic, which is not consistent with a nonuniform spatial distribution if the burst sources are members of a galactic disk or halo population[23,24].

High sensitivity balloon observations[25,26] of bursts with fluence > 10^{-8} erg/cm^2 also suggested that the flattening of the size frequency distribution at low fluence was due to spatial nonuniformities and not due to selection biases. If true, the distribution of low fluence bursts would exhibit anisotropy with a concentration at low galactic latitudes. The BATSE experiment on CGRO was designed to resolve this contradiction.

The CGRO era

Compton's BATSE was designed specifically to investigate weak GRBs and localize them using its eight large 2000 cm^2 sodium iodide scintillation detectors mounted along the eight corners of the spacecraft providing an all-sky coverage. The events recorded are localized to an accuracy of a few degrees using the anisotropic response of the detectors. While this accuracy falls short of that needed to pinpoint visible counterparts, it suffices to map the burster distribution on the sky. In response to a trigger signal from BATSE, the other instruments on Compton[27], viz. Oriented Scintillation Spectroscopy Experiment (OSSE), the Compton Telescope (COMPTEL) and the Energetic Gamma Ray Experiment Telescope (EGRET) switch to a burst data collection mode, if the source occurs within their field of view. While COMPTEL can measure energy spectra of GRBs with high resolution between 0.1 and 10 MeV, it is also capable of localizing sources to about a degree. The EGRET instrument covers the broadest energy range from 20 MeV to 30 GeV. The Compton instruments have detected several unique individual events, the shortest event lasting ~ 100 μs (ref. 28), a few bright GRBs whose emission extends greater than 30 MeV (ref. 29) and two bursts from the region of SGR1900 + 14 (ref. 30). BATSE has also studied the spectra of the brighter bursts and searched for cyclotron line features[31–34]. Even though no definite evidence for lines is found, statistics do not yet support a clear inconsistency between BATSE and earlier mission results.

The most astounding result from BATSE data is that the angular distribution of GRB sources is found to be as isotropic as ever, with no concentration towards the galactic plane, or in directions of any other nearby galaxy or clusters of galaxies[35]. Figure 1 shows the spatial distribution in galactic coordinates of 1000 BATSE bursts[36]. BATSE has also detected fewer weak

bursts[36] (Figure 2) than would be expected for a $-3/2$ power law (shown by dotted lines), implying that the source distribution is not uniform out to infinity. In other words, weaker sources are fewer than expected, indicating thereby a boundedness to the region of source confinement. As burster models have to necessarily account for these important results on spatial and intensity distributions, these results have triggered a flurry of modelling activity, with a raging controversy on the burster distance scale.

Burster models and the distance debate

Prior to the launch of Compton, there existed a wide consensus on the local heliocentric models which favour the galactic neutron star paradigm. In these models the burster population is assumed at a distance of say 100–200 pc but still within the confines of the galactic disk in the solar neighbourhood. One of the early models

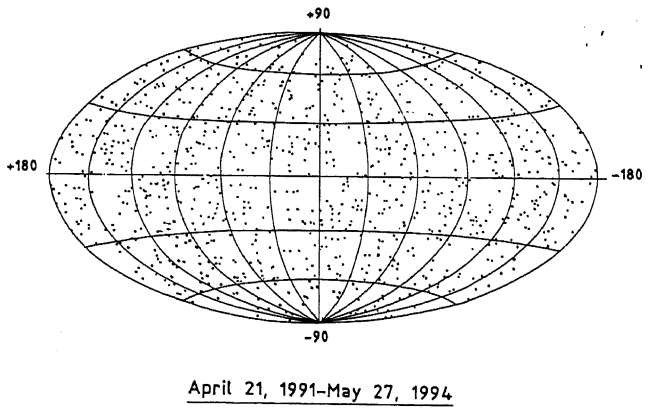


Figure 1. Spatial distribution of 1000 BATSE bursts.

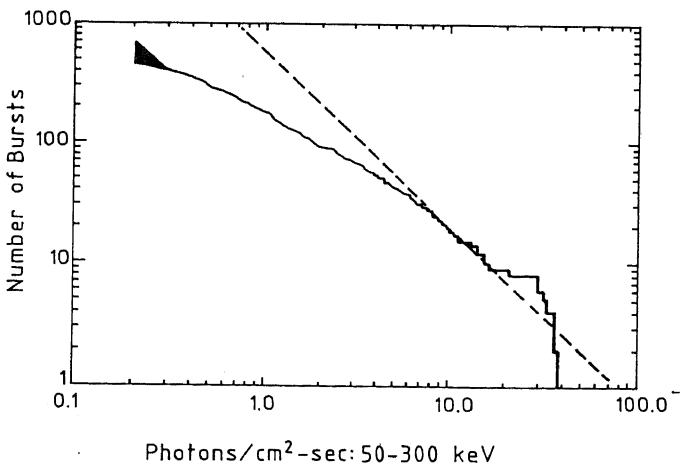


Figure 2. Integral peak flux distribution for 455 BATSE bursts.

focused attention on the retention of a cometary cloud by a collapsing star and subsequent fall/accretion of a solid object of mass $\sim 10^{17-19}$ g onto the neutron star releasing about 10^{36-40} ergs of energy in an explosion[37,38]. Another model, viz. the thermonuclear model, first proposed by Woosley and Taam[39], received more detailed attention over the years and predicts GRBs to occur due to slow accretion of matter from a companion or from the interstellar matter, leading eventually to a helium flash when the critical density and temperature are reached. Gamma radiation is believed to be produced due to magnetic reconnection. Though these models easily explain isotropy, a decrease in the density of sources far away is not expected, as the sun does not occupy a special place in the galaxy and there is no known population of galactic objects at these distances. Considering distances larger than the galaxy's thickness, we run into the opposite problem, i.e. while the radial inhomogeneity is easily explained, an anisotropy towards the galactic plane is not seen.

As of now, with the constraints imposed by BATSE observations on the spatial distribution of faint GRBs, the most widely accepted distance models place the GRB sources either in the outer regions of our galaxy (galactic corona models) or close to the event horizon of our universe (cosmological models). The former model postulates a source population extending about 100 kpc and centered on the galaxy such that the Sun's offset with respect to the galactic centre is negligible. While this galactic corona model explains both isotropy and radial inhomogeneity of burster distribution, the problem here is to populate a halo with neutron stars and find an explanation as to why they became active GRB sources in the halo and not when they were within the galaxy. Population II neutron stars, i.e. those formed by accretion-induced collapse of white dwarfs with birth velocities greater than 1000 km/s are believed to be shot out into the halo surrounding the galaxy[40]. Quakes on neutron stars, which produce about 10^{42} ergs of energy have been suggested as powering mechanisms[41]. As yet, there is no clear evidence supporting the existence of such a structure in the galaxy.

If the cosmological scenario is considered, isotropy follows very naturally, and the flattening of the size frequency distribution is attributed to the effects of the expansion of the Universe[42,43]. The most popular scenario for cosmological bursts involves a collision in a close binary consisting of a black hole and a neutron star or two neutron stars. The orbit of a binary neutron star, particularly those with tight orbits like binary pulsars, decays due to emission of gravitational radiation and eventually coalesces and merges releasing 5×10^{53} ergs mostly as a neutrino burst or as gravitational energy. For GRBs to be detectable at cosmological distances, only a fraction as small as 10^{-3} of this energy is required

to be converted to gamma rays. Such events are estimated to occur in the universe at a rate of 10^{-6} per year per galaxy which is comparable to the GRO burst detection rate. It also follows that other cosmological effects like time dilation, redshift and gravitational lensing should be observable in the weak bursts. Recently an announcement of time dilation in faint BATSE bursts by a factor of 2 has been adduced as evidence[44] for the cosmological distance scale.

At the moment, the mystery of gamma bursters has only deepened and is awaiting some observations or theoretical developments that would settle the dispute on the distance scale. Since neither the location of GRB sources nor their repetition rates are known, pre-planned observations in other wavebands are not feasible. Quiescent counterparts of bursters have been sought both in survey catalogues of unusual objects which might lie in any of the burst error boxes and in deep observations of the more precise error boxes in various wave bands. Since detectors more sensitive than the BATSE are unlikely to be built and flown for some more time to come, the only hope on resolving the GRB puzzle may be pinned down to observations of GRB counterparts at other wavelengths.

Search for counterparts

The success of any search for counterparts relies heavily on the ultimate size of the error box that can be achieved. GRBs being infrequent phenomena, require omnidirectional detectors, as accurate pointing from a single detector is not practicable. The location of the source is determined post facto either by using the angular dependence of several detectors[24,45] or by the triangulation technique[23] using the time difference in the arrival of the signal between two spacecrafts. While the single satellite method is capable of localizing sources with accuracies of only a few degrees, the latter method is necessary to reduce the size of the error box to several tens of arcmin or better so as to facilitate observations at other wavelengths.

The searches for counterparts of the SGRs have been too successful and it now appears certain that SGRs may be related to young magnetized neutron stars in supernova remnants. SGR0526-66, which emitted the historic burst of 5 March 1979, was detected by about ten spacecrafts and triangulation has yielded an association with the SNR N49 in the LMC. This source has burst sixteen times over the years and a faint X-ray source consistent in position with the GRB error box has been detected by ROSAT[46]. The other two repeaters, viz. SGR1806-20 and SGR1900 + 14, have been observed by CGRO to turn ON again[47]. SGR1806-20 has produced 100 bursts irregularly spaced in time between 1979 and 1985 and 6 bursts in 1993 (ref. 47). A galactic SNR G10.0 – 0.3 (ref. 48) and an X-ray source AX1805.7-2025 (ref. 49) have been associated with the SGR1806-20. Simultaneous X-ray and gamma ray bursts have also been observed by the Japanese ASCA satellite and CGRO[50]. SGR1900 + 14 has burst 6 times during 1979–1993 and is probably associated with the galactic SNR G42.8 + 0.6, although its location lies just outside the remnant[51]. At this time, as the nature of classical GRBS is not known, this provides a strong encouragement for counterpart searches as most GRB models require some counterparts to be detectable at some wavelength on some timescale.

Search for counterparts of classical GRBs during the quiescent states, inside GRB errorboxes, has been attempted in all wavebands[52], viz. soft X-ray, ultraviolet, optical, infrared and radio bands. About a dozen optical transient images of astrophysical origin found on archival plates[53] are reported, but their relation to GRBs is unclear. While flaring and fading phase (few hours to few days) searches have technologically limited the observations to X-ray, optical and radio wavebands, until recently searches on timescales shorter than a day were not possible. With BATSE and COMPTEL positions available within a few hours of the burst[54], and the response time of interplanetary networks being about one day[55], fading phase observations are expected to be more fruitful in the near future.

The Interplanetary Network (IPN)

The success of counterpart search depends on the best localizations possible with a network of interplanetary satellites, by the triangulation technique. At the present time, USA's Compton Gamma Ray Observatory and India's SROSS-C2 satellite in near earth orbits, Russia's GRANAT[56] in a highly eccentric orbit around the earth, the international Ulysses probe[57] in a solar polar orbit and the international WIND[58] spacecraft heading for placement at the first Lagrangian point of the Earth–Sun system carry instruments that form the third interplanetary network of GRB monitors.

The GRANAT spacecraft carries three experiments which can record gamma rays, viz. SIGMA – a coded mask telescope; PHEBUS – six Bismuth Germanate scintillation detectors (each 48 cm^2 area) for recording and localizing GRBs; and WATCH – a wide angle telescope. The PHEBUS experiment records bursts of the $> 10^{-6}$ ergs/cm^2 fluence class and the other two experiments compliment this data with their imaging capability. The WIND spacecraft carries two experiments which are capable of recording data on GRBs, viz. the Konus-W gamma ray spectrometer which has two identical sodium iodide detectors (120 cm^2 area) and the Germanium Transient gamma ray spectrometer (TGRS) which is meant for conducting high resolution spectral measure-

ments on GRBs with a sensitivity to fluence greater than a few times 10^{-7} ergs/cm^2.

The hard X-ray detectors on Ulysses consist of two hemispherical thallium-doped caesium iodide shells (40 cm^2 area) each coupled to a photomultiplier tube, providing essentially an all-sky view. The instrument is sensitive to bursts with fluence $> 10^{-6}$ ergs/cm^2 and records a burst once every 3–4 days. The solar probe is presently over the North polar region (June 1995) of the sun.

The SROSS satellites

The Indian SROSS-C and SROSS-C2 satellites carrying the GRB experiment were launched on 20 May 1992 and 4 May 1994 respectively, with the objective of

a) monitoring GRBs in the energy range 20 keV–3 MeV with a sensitivity to bursts of fluence $> 5 \times 10^{-6}$ ergs/cm^2
b) monitor intensity variations with high time resolution (2 ms)
c) search for periodicities in the time histories
d) determine the evolution of the energy spectra with time and
e) search for cyclotron absorption features in the 20–100 keV energy range and look for possible redshifted annihilation radiation line in the 400–500 keV range.

Both satellites were of octagonal prismoidal shape, carrying the payload on their topdeck, and launched into a 45° inclination low earth orbit. The spin stabilized spacecrafts have their spin axis parallel to the detector view axis and along the negative orbit normal (Figure 3), thereby permitting a half sky coverage less the region (~ 34%) occulted by the earth due to low earth orbit. In addition, due to the high inclination orbit, about 30% of observing time is lost owing to passage through the South Atlantic Anomaly (SAA) and high latitude regions.

The experiment on SROSS-C had two sodium doped caesium iodide scintillation detectors, a main (44 cm^2 area), a redundant (11 cm^2 area) and a common microprocessor (CDP1802) based electronics[59,60]. The SROSS-C2 payload (Figure 4), however, carries only the main detector with its associated electronics[61]. The SROSS-C payload could record data on only one event between consecutive readouts, while the payload onboard memory in the case of SROSS-C2 is enhanced to hold the data on seven events. This feature has enhanced the duty cycle of the payload from ~ 40% in SROSS-C to about 70% in SROSS-C2.

SROSS GRBs

The experiment flown on SROSS-C worked very satisfactorily during the entire mission life of 55 days. It recorded 3 candidate GRB events, which were however not detected by GRB detectors onboard other satellites in the interplanetary space and hence remained unconfirmed[62].

SROSS-C2, which has been in orbit for the past one year, has a life expectancy of about 5 years. After the placement of the satellite in its final orbit (620 × 430 km) on 10 July 1994, the GRB payload has been operating flawlessly in meeting with its design goals. The payload records an average of one or two events every month and has recorded 15 events up to 30 May 1995. Table 1 gives the shortlisted SROSS-C2 GRB events. Of these, 13 events have so far been confirmed for detection onboard other GRB monitors. In order to confirm an event, the structure, rise time, decay time, duration and intensity of the temporal profiles recorded on different satellites have to agree with one another within their errors of measurement. Events coincident with Ulysses, BATSE of CGRO, and WIND are marked U, B and W respectively.

The event detected by SROSS-C2 on 6 January 1995 was not detected by Ulysses, presumably due to its faintness. For the event in orbit 1448, Ulysses had a rate increase in its detector, but not a trigger. As the Sun was within the field of view of the detector, this event could possibly be of solar origin. The sensitivity of the SROSS–GRB payload, implies a detection rate of ~ 18 events per year and is consistent with the observations. A mosaic of a few events is shown in

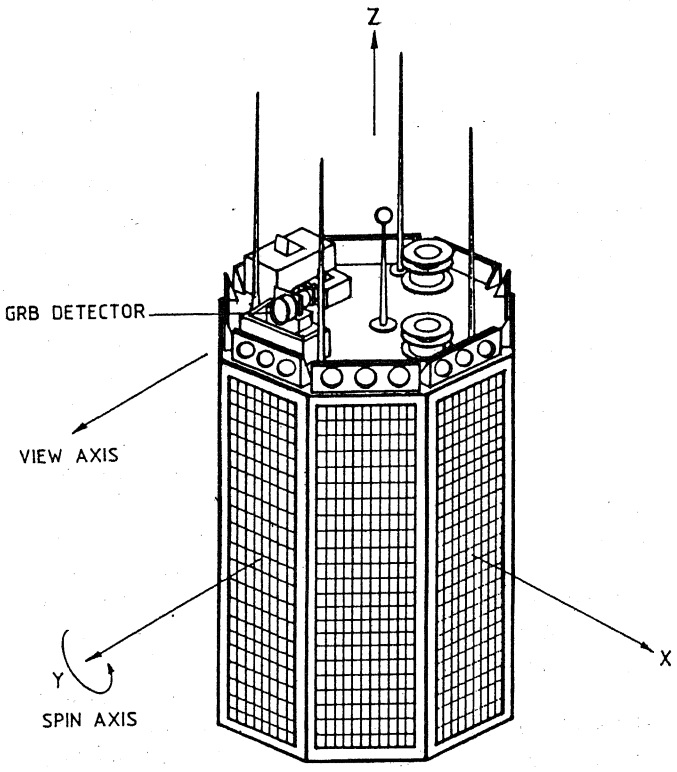


Figure 3. SROSS-C2 with GRB detector mounted on the top deck.

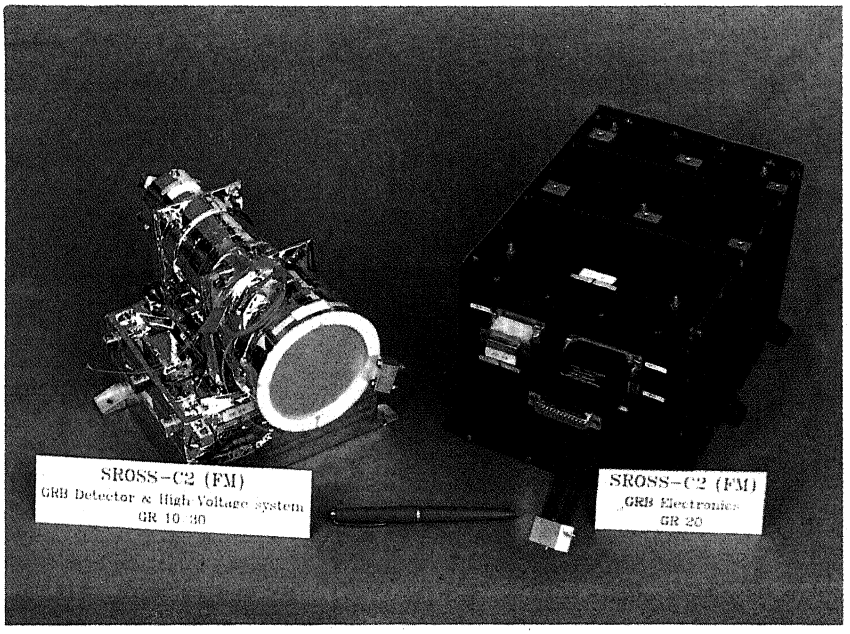


Figure 4. The gamma ray burst payload.

Table 1. SROSS-C2 GRB events

Event orbit	Event date	Time (UT) h m s	Event confirmed by
1126	19 Jul 94	06 00 05.614	U
1296	30 Jul 94	11 01 30.658	U, B
1448	09 Aug 94	11 52 31.791	—
1650	22 Aug 94	20 23 19.878	U
2095	21 Sep 94	05 08 19.368	U, B
2234	30 Sep 94	10 02 06.008	U
2492	17 Oct 94	10 19 59.748	U
3099	26 Nov 94	12 23 05.244	U, B, W
3305	10 Dec 94	03 10 53.906	U, W
3719	06 Jan 95	10 39 57.330	—
4138	03 Feb 95	02 21 29.308	U
4259	11 Feb 95	02 24 57.626	U, B
5101	07 Apr 95	15 15 31.312	U
5497	03 May 95	18 36 10.626	U, B
5622	12 May 95	00 16 48.280	U, W

U, Ulysses; B, BATSE; W, WIND.

Figure 5. As seen from the table, all confirmed events are coincident with Ulysses, while only 5 are coincident with CGRO. The reason probably could be attributed to the fact that Ulysses being a solar probe in an orbit which is presently out of the ecliptic plane has essentially an all-sky view, while both SROSS-C2 and CGRO being low-earth orbiting satellites undergo earth occultation of sky and loss of observing time due to passage through high latitudes and the SAA.

Most events last typically few tens of seconds, with the exception of two, one on 22 August 1994 and the other on 11 February 1995, which lasted about 400 ms and 128 ms respectively. The weakest burst detected by the payload so far has a fluence of about 3×10^{-6} ergs/cm^2. The normal characteristics of GRBs, viz. bimodal distribution of burst durations i.e. < 2 s and > 2 s; spectral hardness and intensity correlation etc. are all seen in our data base, albeit its small size. Analysis of all the confirmed events is being reported elsewhere.

Tracking down the GRB sources

With a few of the recent counterpart observations on Soft Gamma Repeaters being successful, it appears that these might be a separate population of sources, probably related to young pulsars in supernova remnants. However, the nature and origin of 'classical' GRBs remains one of the most compelling astrophysical problems, in view of the isotropy results of BATSE on CGRO. The solution to this problem, at the moment, demands observations that will settle the burster distance scale unambiguously. It is thus evident that a reliable statistical analysis and identification of the phenomenon with a definite class of astronomical objects requires the knowledge of as many source locations as accurately as possible. With each IPN, the triangulation baselines have grown progressively longer and have reached around 6 AU, the longest ever, when Ulysses was close to Jupiter, permitting localizations as fine as 0.1 arcmin^2. Although greater than 2000 GRBs have been recorded till date, about a 1000 or so are localized to a few degrees and only a dozen to few arcmin^2 (ref. 63).

The future efforts on tracking down the sources rely

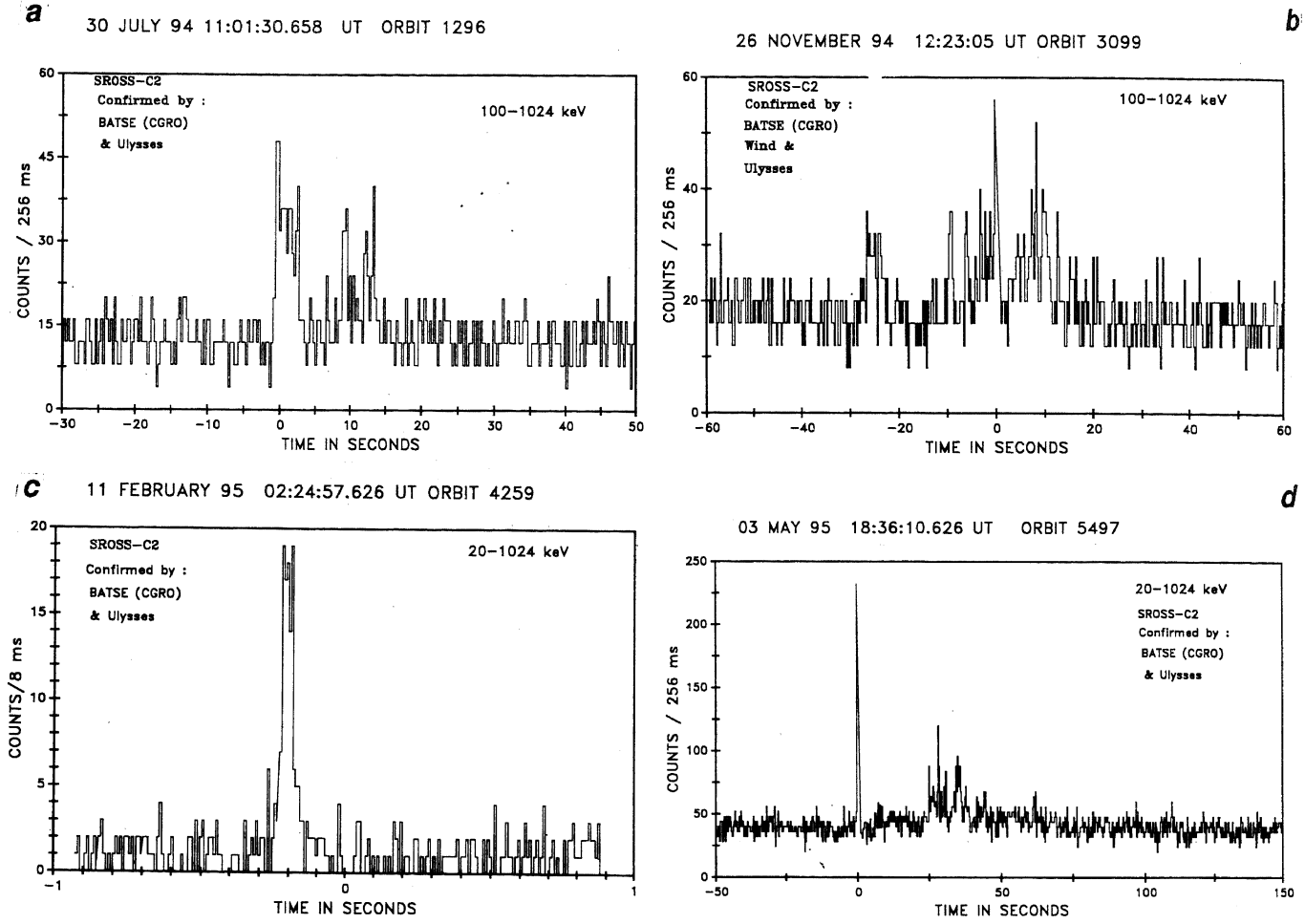


Figure 5. A few events detected by SROSS-C2.

on rapid response campaigns between CGRO and fast slewing ground-based radio and optical telescopes. Multi-wavelength (gamma, X-ray and ultraviolet) observations are planned onboard the high energy transient experiment[64] (HETE) scheduled for launch in late 1995. HETE intends to localize GRBs in the UV band to arcsecond accuracy, although it is not proven that optical flashes accompany GRBs.

Since all near-earth spacecrafts undergo earth occultation, complete sky coverage at all times requires more than one instrument in near-earth space. The GRB experiment on SROSS-C2 is expected to fulfil this function and complement data obtained by CGRO and GRANAT in near-earth space. While the SROSS-C2 payload is not designed to participate in any rapid response campaign, or conduct multiwavelength observations, it is expected to detect about 80–90 bursts during its 5-year life time, of which most are likely to be coincident with Ulysses and WIND. Since Ulysses, WIND and SROSS-C2 do not independently possess the ability to localize GRBs, the class of events, coincident between these three spacecrafts, but not necessarily coincident with CGRO, would form a complement to the CGRO–Ulysses–WIND data base. We hope to record at least a dozen such events. With this, SROSS-C2 would play an important role in the international efforts of tracking down the GRB sources.

1. Klebesadel, R. W., Strong, I. B. and Olson, R. A., *Astrophys. J. Lett.*, 1973, **182**, L85–88.
2. Klebesadel, R. W., Fenimore, E. E., Laros, J. G. and Terrell, J., in *Gamma Ray Transients and Related Astrophysical Phenomena* (eds Lingenfelter, R. E., *et al.*), American Institute of Physics, New York, 1982, pp. 1–15.
3. Fishman, G. J., *et al.*, in Proceedings of the GRO Science Workshop, NASA (ed. Johnson, W. N.), 1989, pp. 2–39.
4. BATSE news in *Sci. Rep.*, Compton Observatory, 1995, June 8, p. 182.
5. Hurley, K., in *Cosmic Gamma Rays, Neutrinos and Related Astrophysics* (eds Shapiro, M. and Wefel, E.), Kluwer, Boston, 1989, p. 337.
6. Higdon, J. C. and Lingenfelter, R. E., *ARA&A*, 1990, **28**, 401–436.
7. AIP Conference proceedings No. 307 on Gamma Ray Bursts (eds Fishman G. J., Brainerd, J. J. and Hurley, K.), 1994.
8. Cline, T. L., *et al.*, *Astrophys. J.*, 1982, **255**, L45–48.
9. Mazets, E. P., *et al.*, *Nature*, 1979, **282**, 587–589.
10. Barat, C., *et al.*, *Astron. Astrophys.*, 1979, **79**, L24–25.
11. Terrell, J., *et al.*, *Nature*, 1980, **285**, 383–385.
12. Mazets, E. P., *et al.*, *Nature*, 1981, **290**, 378–382.

13. Mazets, E. P., *et al.*, *Astrophys. Space Sci.*, 1982, **82**, 261–282.
14. Murakami, T., *Nature*, 1988, **335**, 234–235.
15. Liang, E. P., *Astrophys. J.*, 1986, **304**, 682–687.
16. Helfand, D. J. and Long, K. S., *Nature*, 1979, **282**, 589–591.
17. Pizzhichini, G., *et al.*, *Astrophys. J.*, 1986, **301**, 641–649.
18. Pederson, H., *et al.*, *Astrophys. J. Lett.*, 1983, **270**, L43–47.
19. Barat, C., *et al.*, *Astrophys. J.*, 1984, **280**, 150–153.
20. Laros, J. G., *et al.*, *Astrophys. J.*, 1985, **290**, 728–734.
21. Motch, C. *et al.*, *Astron. Astrophys.*, 1985, **145**, 201–205.
22. Hartmann, D. H. and Pogge, R. W., *Astrophys. J.*, 1987, **318**, 363–369.
23. Atteia, J. L., *et al.*, *Astrophys. J. Suppl.*, 1987, **64**, 305–382.
24. Mazets, E. P., *et al.*, *Astrophys. Space Sci.*, 1981, **80**, 3–143.
25. Beurle, K., *et al.*, *Astrophys. Space Sci.*, 1981, **77**, 201–214.
26. Meegan, C. A., *et al.*, *Astrophys. J.*, 1985, **291**, 479–485.
27. Proceedings of the GRO Science Workshop, NASA (ed. Johnson, W. N.), 1989.
28. Fishman, G. J., *et al.*, in AIP Conference Proceedings (eds Paciesas, W. S. and Fishman, G. J.), 1992, p. 13.
29. Hurley, K., *et al.*, *Nature*, 1994, **372**, 652.
30. Koveliotou, C., IAU Circ. 5567 & 5592, 1992.
31. Teegarden, B. J., *et al.*, *BAAS*, 1991, **23**, 1470.
32. Palmer, D. M., *et al.*, *BAAS*, 1992, **24**, 1259.
33. Schaefer, B. E., *et al.*, in AIP Conference Proceedings (eds Paciesas, W. S. and Fishman, G. J.), 1992, p. 180.
34. Band, D., *et al.*, *BAAS*, 1992, **24**, 1258.
35. Meegan, C. A., *et al.*, *Nature*, 1992, **355**, 143.
36. Meegan, C. A., *et al.*, in AIP Conference Proceedings No 307 (eds Fishman G. J. *et al.*), 1994, p. 3.
37. Harwit, M. and Salpeter, E., *Astrophys. J. Lett.*, 1973, **186**, 37.
38. Schklovskii, I., *Soviet Astronomy*, 1974, **51**, 665.
39. Woosley, S. and Taam, R., *Nature*, 1976, **263**, 101.
40. Li, H. and Dermer, C. D., *Nature*, 1992, **359**, 514.
41. Pacini, F. and Ruderman, M, *Nature*, 1974, **251**, 399.
42. Paczynski, B., *Astrophys. J. Lett.*, 1986, **308**, L43.
43. Narayan, R., Paczynski, B. and Piran, T., *Astrophys. J. Lett.*, 1992, **395**, L83.
44. Norris, J. P., *et al.*, *Astrophys. J.*, 1994, **424**, 540–545.
45. Cline, T. L., *et al.*, *Astrophys. J. Lett.*, 1984, **286**, L15–18.
46. Rothschild, R. E., *et al.*, *Nature*, 1994, **368**, 432.
47. Koveliotou, C., *et al.*, *Nature*, 1994, **368**, 125.
48. Kulkarni, S. R., *Nature*, 1993, **365**, 33.
49. Cooke, B. A., *Nature*, 1993, **366**, 413.
50. Murakami, T., *et al.*, *Nature*, 1994, **368**, 127.
51. Hurley, K. *et al.*, *Astrophys. J.*, 1994, **431**, L31.
52. Shaefer, B. E., in AIP Conference Proceedings No. 307 (eds Fishman, G. J., *et al.*), 1994, p. 382.
53. Hudec, R., *Astron. Lett. Comm.*, 1993, **28**, 359.
54. Kippen, R. M., *et al.*, in AIP Conference Proceedings No. 307 (eds Fishman, G. J. *et al.*), 1994, p. 418.
55. Hurley, K., in AIP Conference Proceedings No. 280 (eds Friedlander, M., Gehrels, N. and Macomb, D. J.), 1993, p. 769.
56. Barat C., *A&A Suppl.*, 1993, **97**, 43–48.
57. Hurley, K., *A&A Suppl.*, 1992, **92**, 401–410.
58. Owens, A., *et al.*, *IEEE-NS*, 1991, **38**, 559–567.
59. Marar, T. M. K., *et al*, *A&A*, 1994, **283**, 698.
60. Ramakrishna Sharma, M. R., *et al.*, *IEEE NS*, 1993, **40**, 1989–1997.
61. Marar, T. M. K., *et al.*, *J. Spacecraft Technol.*, 1995, **5**, 75.
62. Kasturirangan, K., *et al.*, *A&A*, 1994, **283**, 435.
63. Hurley, K., in AIP Conference Proceedings No. 307 (eds. Fishman G. J. *et al.*), 1994, p. 359.
64. Ricker, G., *et al.*, in *Gamma Ray Bursts*, Cambridge University Press, 1992, p. 288.

ACKNOWLEDGEMENTS. We thank the staff of the Technical Physics Division for their unstinted efforts during the payload development. Thanks are also due to the SROSS project team and the staff of the satellite tracking centre (ISTRAC) for their continuing support in the payload operations.

Received 12 August 1995; accepted 27 September 1995

Quantum Monte Carlo techniques: Chemical applications

Charusita Chakravarty

Department of Chemistry, Indian Institute of Technology, New Delhi 110 016, India

Quantum Monte Carlo (QMC) methods are stochastic simulation methods for quantum many-body systems. Here we provide an overview of QMC methods designed for simulation of atomic and molecular systems, and discuss main ideas behind the variational, diffusion and path integral QMC methods. We also review application of these techniques to some problems of interest in chemistry.

> The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are (thus) completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be solvable. It therefore becomes desirable that approximate practical methods of applying quantum mechanics should be developed, which can lead to an explanation of the main features of complex atomic systems without too much computation.

SINCE P. A. M. Dirac wrote the above lines in 1929 (ref. 1), both the range of approximate quantum mechanical methods, and the notion of what constitutes 'too much computation' have undergone dramatic changes.

Over the past few decades, chemists have come to understand the interactions and dynamics of small molecules and clusters very well. Both theory and experiment can, for instance, provide a very accurate

understanding of the H_2O molecule and a reasonable one of the dimeric and trimeric aggregates $(H_2O)_2$ and $(H_2O)_3$ (refs 2,3). The macroscopic end of the scale, such as the thermodynamics and spectroscopy of bulk liquids and solids is again well understood from a phenomenological point of view; for example, bulk water is known to be a complex fluid with an elaborate phase diagram[4]. The conceptual connection between microscopic and macroscopic behaviour is made by statistical mechanics. Computer simulation methods complement the elegant but simplified statistical mechanical models by providing a much clearer connection between experimental and calculated quantities, the ability to make quantitative or semiquantitative predictions and the opportunity for performing numerical or computer experiments for systems or conditions which may not be easily realizable in a laboratory.

Currently, most simulations are performed using either classical Monte Carlo methods or the molecular dynamics approach[5]. Applications of these methods range from geochemistry to protein structure[6,7]. Both techniques assume the Born–Oppenheimer separation of electronic and nuclear motion. The nuclei are assumed to be classical particles moving according to the Newtonian equations of motion on the electronic potential energy surface (PES). The PES itself is constructed as a sum of pair (or at most three- and four-body) interactions between atoms, i.e. the total interaction energy of a system is assumed to be the sum of the interaction energies between all unique pairs. These few-body potentials are derived from independent studies of small aggregates and provide the necessary simulation input about the details of the microscopic interaction[8]. Classical Monte Carlo techniques generate the equilibrium statistical mechanical properties of a system, usually within the canonical ensemble, such as the average potential energy, structural quantities such as the radial distribution function and phase transition behaviour. For an N-atom system, such Monte Carlo techniques construct a random walk through the $3N$-dimensional configuration space with a bias such that for sufficiently long walks, a set of configurations, $\mathbf{x}$, distributed according to the Boltzmann formula $e^{-\beta V(x)}$ is generated. If N is sufficiently large and suitable boundary conditions are applied, then equilibrium properties of the bulk system can be generated. MC integration is an inherently multidimensional approach with a statistical accuracy in the results that is proportional to $(n/\tau)^{1/2}$, where n is the number of MC configurations and τ is a measure of the degree of correlation between successive configurations. While τ does tend to increase with dimensionality of the system, this dependence is generally small. Consequently, MC methods scale relatively favourably with system size, in terms of computational cost.

Classical molecular dynamics, instead of using a random walk approach, involves numerical integration of the classical equations of motion to generate the trajectory of the N-particle system through phase space. Provided certain conditions are satisfied (ergodicity, sufficient length of trajectory and the like), the phase space points sampled during the course of an MD run provide information on both the equilibrium and dynamical statistical mechanical quantities. Thus, unlike MC, MD methods can provide time-correlation functions related to transport properties and spectroscopy. However, MD can prove inefficient relative to MC for systems with multiple time scales. For example, in binary solid solutions, the rates of diffusion of the two species are much slower than the high frequency lattice vibrations[9].

An important development in the past decade has been the formulation of *ab initio* MD methods, which attempt to overcome the problems of inadequate parametrization of the PES surface. Such problems are most acute in the case of semiconductor and metal systems; the delocalized nature of the valence electrons in such materials implies that the pair potential approximation is very poor. This is also true for chemical reactions where changes in electronic structure and nuclear configuration are very strongly correlated. Polarizability effects are also notoriously hard to incorporate within a pair potential approach. The *ab initio* MD methods attempt to solve this problem by solving the electronic structure problem simultaneously with the integration of the classical equations of motion and, in principle, require no input other than electronic and nuclear masses and electric charges.

The most successful *ab initio* MD approach to date is the Car–Parinello (CP) formulation. The CP technique uses the Born–Oppenheimer approach to separate motion of the ionic cores from that of the valence electrons. The motion of ionic cores is treated by using classical MD equations of motion. To obtain a suitable electronic PES, the ground state electronic energy for any given ionic configuration must be found. To do this both accurately and efficiently, three important simplifications are made: (i) Density functional theory (DFT) is used to model electron–electron interactions. DFT is a mean-field approach which introduces electron correlation effects in an approximate way. Results for a variety of ground state properties using DFT have been shown to be accurate to within a few per cent. (ii) Pseudopotential theory is used to model electron–ion interactions. Pseudopotentials smooth out the strongly varying regions of the potential for small ion–electron distances while retaining the salient characteristics of valence electron in a bulk medium. (iii) Dynamical minimization techniques are based on the insight that the variational solution to the density functional equations can be obtained by a constrained optimization scheme in which basis functions for the electronic structure problem are

treated as dynamical variables. The advantage of this approach is that the real-time dynamics of the ionic cores and the fictitious dynamics associated with the electronic variables can be coupled together in a very powerful computational scheme. The net result is that the classical motion of atoms on the ground state Born–Oppenheimer PES is obtained. The CP method has provided remarkable results for many problems of current interest in materials science[10]. The technique is at least an order of magnitude more expensive than classical MD.

Quantum Monte Carlo (QMC) methods, as the name suggests, are stochastic simulation methods for quantum many-body systems[11–16]. The term QMC covers a battery of techniques used in a variety of fields such as condensed matter physics, nuclear physics, statistical mechanics and chemistry. This review is intended to provide an overview of QMC methods designed for simulation of atomic and molecular systems of interest in chemistry.

One situation in which QMC methods become essential for chemical systems is when the atomic nuclei are sufficiently light or temperatures are sufficiently low that the approximation of treating nuclear motion classically becomes a poor one. For example, quantum effects are important in the water and ice because of hydrogen-bonding. Electron and proton transfer processes, vital to many biochemical processes, are intrinsically quantum mechanical. Many solids such as ammonium halides and solid H_2 show phase transitions, with large isotope effects indicative of strong quantum effects[4]. Even solid argon must be modelled with significant quantum corrections to thermodynamic properties[5].

Electronic structure problems are a second area of chemistry where QMC methods are applicable when accuracy requires improvements over density functional theory (the basis of Car–Parrinello) or when the Born–Oppenheimer approximation fails due to strong coupling between nuclear and electronic motion[13].

The review is organized as follows. Sections 1, 2 and 3 present the essential ideas behind the variational, diffusion and path integral Monte Carlo methods respectively. Section 4 briefly reviews the application of QMC methods to some problems of interest in traditional chemistry where the simulations provide interesting insights and connections between experimental data and simplified theoretical models.

1. Variational quantum Monte Carlo methods

The variational Monte Carlo (VMC) approach is based on the variation principle which states that for a given Hamiltonian $\hat{H}$, the energy E_t associated with a trial wavefunction ψ_t, defined as

$$E_t = \int \psi_t^* \hat{H} \psi_t \, d\tau / \int \psi_t^* \psi_t \, d\tau \qquad (1)$$

will always be greater than or equal to the exact ground state energy E_0 of the system[12]. Minimizing E_t with respect to the parameters of ψ_t provides the best approximation to the ground state wavefunction within the limitations of a given functional form. Conventional quantum chemical methods expand the trial wavefunction ψ_t as a linear combination of a suitable set of basis functions and then utilize the variation theorem to obtain the optimum set of linear combinations. While such methods can provide very accurate energies and wavefunctions, they scale very inefficiently with the system size or dimensionality. For example, configuration interaction methods scale as N^7, where N is the number of basis functions, implying that such studies are only feasible for small molecules.

VMC methods use the favourable scaling of Monte Carlo methods with system size by interpreting $\psi_t^* \psi_t$ as a probability distribution. Standard random walk methods can be used to create a set of configurations, $\mathbf{x}$, distributed according to $\psi_t^*(\mathbf{x}) \psi_t(\mathbf{x})$ and E_t can be defined as the average over this set of configurations of the local energy, $E_t = \hat{H} \psi_t(\mathbf{x})/\psi_t(\mathbf{x})$. The form of the trial wavefunction is usually derived from a knowledge of the basic physics of the problem and the parameters can then be optimized by minimizing E_t or, better still, the variance in E_t.

VMC methods are computationally relatively inexpensive and can provide considerable physical insight. The limitations of VMC methods are two-fold: (i) the accuracy of the calculation is limited by the choice of functional form of ψ_t; (ii) being essentially a ground state method, information on excited state energies and wavefunctions is difficult to obtain though not impossible. Despite these limitations, VMC calculations are capable of providing accurate results for a variety of systems. As an example of an application to solid state electronic structure calculation, the bulk cohesive energy of diamond and graphite was calculated to within 1% of the experimental value using VMC[17]. VMC simulations of helium have provided very accurate results, when compared with experiment, for properties of helium in the bulk, at surfaces and in droplets[18].

2. Diffusion Monte Carlo methods

Diffusion Monte Carlo (DMC), like VMC, is a zero-temperature approach that is geared towards finding the ground state wavefunction. Unlike VMC, which provides an upper bound to the ground state energy, DMC is, in principle, capable of generating the *exact* ground state energy and wavefunction[11–16].

The starting point of the DMC method is the time-dependent Schrödinger equation

$$i\hbar \frac{\partial \Psi}{\partial t} = \hat{H} \Psi , \qquad (2)$$

where t is the time coordinate. If t is treated as an imaginary quantity $t=-iu$ (where u is real), then the above equation can be rewritten as:

$$-\hbar \frac{\partial \Psi}{\partial u} = \hat{H}\Psi, \tag{3}$$

or

$$\Psi_{\text{final}} = e^{-u\hat{H}/\hbar}\Psi_{\text{initial}}. \tag{4}$$

By an analogy with the density operator, $e^{-\beta\hat{H}}$, used in statistical mechanics, imaginary time in the above equation can be interpreted as proportional to the inverse temperature, $\beta = 1/k_B T$ such that $u = \beta\hbar$, T is the absolute temperature and k_B is the Boltzmann constant. It can be shown that regardless of the choice of Ψ_{initial}, Ψ_{final} will converge exponentially to the ground state wavefunction Ψ_0 with energy E_0 in the limit $\beta \to \infty$, $T \to 0$. If it is assumed that the ground state wavefunction is nodeless, as is the case for bosons or distinguishable particles, then the above equation can be interpreted as the evolution of an initial, arbitrary probability distribution to the final stationary state distribution given by Ψ_0. To understand the basis of the computational scheme, consider the evolution of the distribution over a very small imaginary time interval, $\varepsilon = \beta/M$, from time u_i to u_{i+1}. Then it is possible to factorize $e^{-\beta\hat{H}/M}$ as $e^{-\beta\hat{K}/M}$ $e^{-\beta\hat{V}/M}$, where $\hat{K}$ and $\hat{V}$ are the kinetic and potential energy operators respectively. Considering a one-dimensional system for notational simplicity, the distribution at time u_i can be represented as a set of points distributed on the x-axis according to the initial distribution $\Phi_i(u_i, x_i)$. To carry out the evolution over the time interval ε, we first apply $e^{-\beta\hat{V}/M}$ on $\Phi_i(u_i, x)$ to give $e^{-\beta(V(x)-E_T)/M}\,\Phi_i(u_i, x)$, where $V(x)$ is measured with respect to some energy E_T. This will imply that in regions of low potential energy, the probability distribution will be enhanced whereas in regions of high potential energy, the distribution will be attenuated (birth/death process). To complete the evolution we must then act using the operator $e^{-\beta\hat{K}/M}$ which results in

$$\sqrt{m/2\pi\varepsilon}\; e^{-m/2\varepsilon(x_{i+1}-x_i)^2}\; e^{-\beta V(x_i)/M}\; \Phi_i(u_i, x_i). \tag{5}$$

Note that the positions obtained at the end of the time step, x_{i+1}, are Gaussian distributed about the positions x_i at time u_i. This can be interpreted as a diffusion step with diffusion constant $\hbar/2m$. Thus a short-time propagation corresponds to a birth/death step followed by a diffusion step. Repeated application of such short time evolutionary operators will result in convergence to the ground state.

For all but the simplest system, the naive DMC algorithm described above must be made more efficient by introducing importance sampling. Importance sampling biases the random walk towards the most important regions of configuration space by using the information provided by a good approximate wavefunction, usually derived from a VMC calculation. A different formulation of this zero-temperature approach, equivalent to the DMC method, is Green's Function Monte Carlo (GFMC).

Monte Carlo methods require the definition of a probability distribution function that is always positive. In the case of VMC simulations, the modulus squared of the wavefunction provides a positive probability distribution function. In DMC simulations of bosons or distinguishable particles, the nodeless ground state wavefunction can be treated as a probability distribution. In the case of interacting fermions, however, the ground state cannot be assumed to be nodeless. The anti-symmetrization requirement for fermions implies that nodes will exist in the wavefunction whenever positions of two fermions with identical spins coincide. Continuity of the wavefunction and its first derivative then implies that the wavefunction must change sign at these nodal planes. Consequently, the interpretation of the ground state wavefunction as a probability distribution is invalid, leading to the well-known 'fermion-sign' problem. Fixed-node methods which are surprisingly accurate though not, in principle, exact, have been developed to circumvent this problem. A very important application of this fixed-node DMC technique has been in deducing the exchange correlation functional for the uniform electron gas, which forms the basis of most density functional and *ab initio* MD calculations[19].

3. Path integral Monte Carlo method

Path integral Monte Carlo (PIMC) techniques allow for simulations of quantum many-body systems in the canonical ensemble (constant number N, volume V and temperature T). The equilibrium properties of a system in the NVT ensemble are determined by the canonical partition function, Q. For a quantum system, Q must be expressed as a trace over the density operator $e^{-\beta\hat{H}}$; evaluation of the trace in the coordinate representation then leads to:

$$Q = \text{Tr}\{e^{-\beta\hat{H}}\} = \int d\mathbf{x}\, \langle \mathbf{x}|e^{-\beta\hat{H}}|\mathbf{x}\rangle, \tag{6}$$

where $\mathbf{x}$ is the position vector for the N-particle system. Path integral methods interpret $e^{-\beta\hat{H}}$ as a propagator in imaginary time (note the similarity with DMC) and express the density matrix elements as the functional integral

$$\langle \mathbf{x}|e^{-\beta\hat{H}}|\mathbf{x}'\rangle = \int \mathcal{D}(\mathbf{x}(u))e^{-S(\mathbf{x}(u))/\hbar}, \tag{7}$$

where the integral represents the sum over all possible paths, $\mathbf{x}(u)$, in imaginary time u which satisfy the condition that $\mathbf{x}(0) = \mathbf{x}$ and $\mathbf{x}(\beta\hbar) = \mathbf{x}$. The weighting of individual paths is given by the exponential of the

action in imaginary time, $S(\mathbf{x}(u))$. Note that only cyclic quantum paths with the same initial and final point can contribute to Q. The classical path is the path of least action[20].

PIMC methods involve finding paths which contribute the most to the partition function. For example, in the discretized path integral approach, any quantum path is represented as a set of configurations, $\{\mathbf{x}_i\}$, at equispaced intervals in imaginary time, $u = 0$, $(\beta\hbar/M)$, $(2\beta\hbar/M)$, ..., $\beta\hbar$. Using this representation, Q for an N-particle system can be written as

$$Q = \frac{1}{N!}\left(\frac{Mm}{2\pi\beta\hbar^2}\right)\int d\mathbf{x}_1\, d\mathbf{x}_2 \ldots d\mathbf{x}_M \times \exp\left\{\frac{-Mm}{2\beta\hbar^2}\,(x_{12}^2 + x_{23}^2 + \ldots x_{M1}^2)\right\} \times \exp\left\{\frac{-\beta}{M}\,(V(\mathbf{x}_1) + V(\mathbf{x}_2) + \ldots V(\mathbf{x}_M))\right\}, \quad (8)$$

where $\mathbf{x}_{ij} = |\mathbf{x}_i - \mathbf{x}_j|$. The partition function in equation (8) has the same structure as the classical configurational integral with $3NM$ spatial degrees of freedom. The interpretation of the above partition function is, however, somewhat novel. In the classical limit, a particle is characterized by 3 position degrees of freedom. In the discretized path representation, the single quantum particle is replaced by a cyclic polymer of M subunits, with adjacent subunits held by a harmonic interaction; the external potential felt by each subunit is $V(\mathbf{x}_i)/M$. The average extent of the polymer is proportional to the thermal de Broglie wavelength of the particle. Equation (8) can be used directly in conjunction with the Metropolis algorithm to simulate a many-particle quantum system. It is, however, often computationally more efficient to employ alternative representations of the path integral[21]. The common feature of all such PIMC methods is the quantum classical isomorphism – by introducing auxilliary degrees of freedom (such as the $M-1$ additional position degrees of freedom for the polymer subunits), the trace over the density operator is written as a classical configurational integral of much higher dimensionality. The fermion sign problem present in DMC recurs also for the finite temperature formulation and methods analogous to the zero-temperature are currently being devised. An interesting and unsolved problem is whether the imaginary time PIMC formulation can be adapted for real-time problems. While this is one approach towards achieving the tantalizing goal of devising a simulation method for many-body quantum dynamics, so far it has met with little success[21].

PIMC methods have been applied extensively to molecular systems. The best studied system is liquid ^{4}He which shows a Bose–Einstein transition to a superfluid state; simulation results compare very well to the experiment[12]. Good approximation methods for finite temperature fermionic systems have been developed recently and applied to liquid ^{3}He and electronic structure[22,23]. An interesting possibility is combining PIMC methods for nuclei with Car–Parinello techniques for obtaining the electronic structure. While prohibitively expensive at present for bulk systems, it has been applied to CH_5^+. The structure of CH_5^+ has long been controversial with arguments that it is a unique carbonium atom as opposed to the theory that it is the prototype for a nonclassical hyper-coordinate carbocation, with a ground state dominated by 3-centre-2-electron bonding. The combined PIMC and *ab initio* MD calculations indicate that the quantum ground state is dominated by 3*c*-2*e* bonding[24,25].

4. Some applications

Surface adsorption

The importance of surface adsorption, from practical applications in heterogeneous catalysis to theoretical models for phase transitions in two-dimensional systems, does not require emphasis. As a result, classical simulation of surfaces and interfaces is by now a well-developed field[26]. Appreciable quantum effects can, however, be expected for light adsorbates, e.g. H_2, He, Ne, N_2 and CO, specially at low temperatures. One of the first surface adsorption problems to be studied by PIMC was ^{3}He and ^{4}He adsorption on graphite[27]. The quantum simulations corroborated experimental results for the different phase diagrams of the two isotopes and provided interesting insights into the microstructure associated with different phases. In addition, DMC/GFMC calculations have been carried out to explore the zero temperature limit of the He/graphite system[12]. In contrast to atomic substrates, diatomic molecules can exhibit a range of orientational transitions. For example, N_2/graphite shows a herringbone transition; PIMC simulations have quantified that the lowering of the transition temperatures due to quantum effects is 10% (30 K) (ref. 28). The CO/graphite system is analogous to the N_2 adsorption problem except for a heat capacity anomaly at approximately 5 K. Recent PIMC simulations have shown that this must correspond to the head-to-tail ordering transition for the two-dimensional CO system[29]. To my knowledge, there have been no zero-temperature studies of molecular adsorption, largely because interest in this area centres on finite temperature phase transitions.

Solvated electrons

The solvated electron; a 'free' electron trapped in a cavity formed by solvent molecules, plays an important role in a variety of chemical phenomena. The hydrated

electron has been shown to be a transient species of crucial importance in solution photochemistry since its identification more than 30 years ago[30,31]. The solvated electron is central to the metal–insulator transitions observed in alkali metal–ammonia solutions and related systems. The localized versus extended states responsible for the metal–insulator transition are also connected to electron mobilities in fluids and other disordered media; therefore the general problem of excess electron mobilities in a variety of solvents has attracted much theoretical and experimental attention[32–34]. The structure of a solvated electron in a fluid will be determined by the competition between two factors: a tendency to delocalize and occupy as large a volume as possible in order to minimize the kinetic energy and a tendency to minimize potential energy by optimizing electron–solvent interactions. In the case of the dominant contribution to the electron–solvent interaction being an excluded volume or repulsion interaction, the electron will tend to localize in regions of low fluid density. If the attractive interactions can be enhanced in these low density traps (for example, by the presence of oriented solvent dipoles) then the size of the cavity can be substantially reduced. The fact that the characteristics of solvated electron systems will depend on this type of complex interplay between a variety of factors, makes simulations a virtual necessity. Different simulation techniques can then be used to provide complementary information on these systems.

Among the best-studied solvated electron systems, both from an experimental and simulation perspective, are the alkali metal–ammonia solutions which have a long experimental history[35]. Alkali metals dissolve in liquid ammonia and, at very low concentrations, form a light blue solution. Spectroscopic and other evidences indicate that, as in the case of F-centres in alkali halides, the blue colour is due to the optical spectrum of an isolated solvated electron. As the concentration of metal is increased, the magnetic susceptibility falls, indicating that some type of spin pairing process is in operation. However, as the metal concentration increases, a metal insulator transition takes place and a bronze-coloured liquid metal is formed. At still higher concentrations, metallic solid compounds are formed. The dilute limit is ideally treated by PIMC methods[36]. The results show the trapping or localization of single electrons in solvent cavities with an effective size of about 4 Å. In fact, a transition from a delocalized to a localized, trapped state is seen as solvent concentration is increased. The results of the simulation agree well with experiment as well as with predictions of the RISM-polaron theory[37]. The metallic limit with high electron concentration can be treated by Car–Parrinello methods[38–40]. The intermediate regime, especially where spin-pairing occurs, can be treated by both approaches. In the spin-pairing regime, the simulations show that electron-pairs form singlet states which occupy a peanut-shaped cavity with the two peaks in the electron density separated by 7 Å. The triplet state is unstable with respect to the singlet state by approximately 0.6 eV. Thus, this series of simulations on the alkali metal–ammonia solutions has provided considerable insight into the microstructure and dynamics associated with the metal–insulator transition.

Quantum clusters

Clusters have attracted much attention by virtue of being small systems which bridge the transition from molecular to bulk properties[41–44]. While clusters can show properties similar to those of bulk matter, e.g. collective behaviour analogous to a phase transition, they are sufficiently small to be studied at the same level of microscopic detail as molecular systems. In this section, some of the applications of QMC methods to quantum clusters are summarized. Quantum clusters denote small, finite atomic or molecular systems (about 10–1000 atoms) where the nuclei have appreciable quantum character. Recent experiments on rare-gas–SF_6 clusters have added to the interest in quantum clusters[45]. Electronic structure aspects, of special relevance to metallic and semiconductor clusters, are not considered here. It should be pointed out that cluster simulations have played an important role in the development of QMC methods; for example, one of the first applications of VMC and GFMC methods was to helium clusters to test the reliability of liquid drop models for clusters[45].

Atomic clusters, such as those of He, Ne, H_2 and D_2 have been very useful in understanding quantum effects on phase transitions[46–54]. For example, ^4He_n with $n > 60$ shows clear remnants of the bulk superfluid transition and QMC simulations can provide a corresponding microscopic model[53]. Bulk melting has its cluster analogue and has been studied in clusters with a view to understanding the role played by quantum fluctuations in modifying phase transition characteristics. For example, Figure 1 shows the lowering of the position and height of the specific heat curve for a Ne_{13} cluster due to quantum effects[48]. Increasing quantum delocalization can be shown to result in a cluster solid–liquid transition analogous to the thermal melting transition[55]. The effects of such delocalization can be seen in Figure 2, where the angular distributions for the Ne_{13}, $(para\text{-}H_2)_{13}$ and $(ortho\text{-}D_2)_{13}$ are compared with those of a classical thirteen atom Lennard–Jones cluster[49]. A purely quantum analogue of the classical binary phase separation can be shown to occur in isotopically mixed clusters using PIMC methods[56]. Figure 3 shows the density profiles for the two isotopic species in a mixed *para*-H_2/*ortho*-D_2 cluster; the heavier isotope can be seen to be preferentially located in the cluster interior.

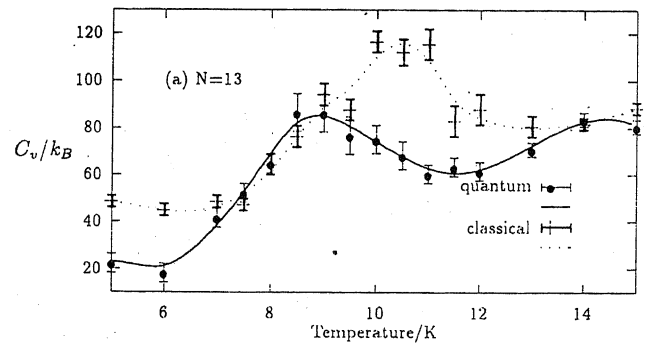


Figure 1. The specific heat, C_v/k_B, as a function of temperature for a Ne_{13} cluster. Results from path integral quantum Monte Carlo and classical simulations are compared.

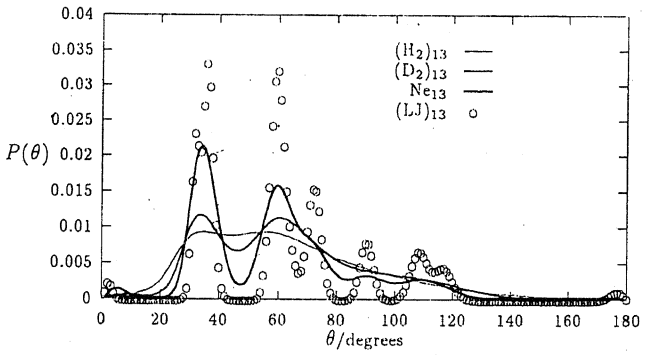


Figure 2. Angular distributions, $P(\theta)$, for thirteen atom clusters at 5 K. The distribution for the classical Lennard–Jones cluster, $(LJ)_{13}$, is compared with that for quantum clusters of Neon, *para*-H_2 and *ortho*-D_2. The latter two species can be treated as pseudo-atomic systems interacting via a spherically symmetric pair potential at low temperatures. The peak at 60° is characteristic of the icosahedral symmetry of small Lennard–Jones clusters.

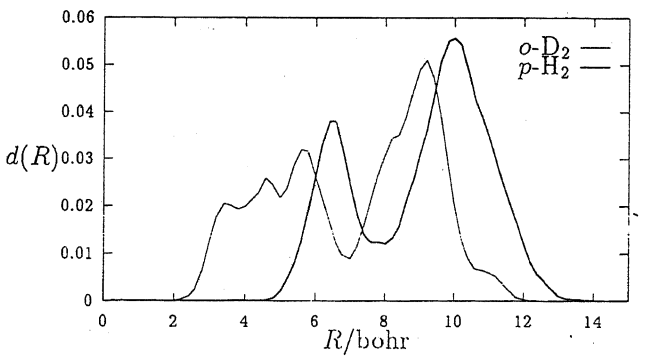


Figure 3. Radial density profiles, $d(R)$ for the two isotopic species in the $(P\text{-}H_2)_9(O\text{-}D_2)_9$ cluster at 2.5 K. The radial density profile is a measure of the probability of finding a particle of a particular species at a distance R from the centre of mass of the cluster.

A second category of clusters with large quantum effects are the so-called solvent clusters such as $(H_2O)_n$, $(NH_3)_n$ and $(HF)_n$. Understanding the structure and dynamics of such systems is of particular importance from the point of view of solvation, liquid structure and hydrogen bonding. Small aggregates of two, three or four monomers have been intensively studied by QMC methods (mostly DMC). Such studies are very valuable from the point of view of deriving accurate intermolecular potentials by comparison and fitting to spectroscopic data.

Electron transfer

Electron transfer is fundamental to a variety of chemical and biochemical processes. An electron transfer process may be represented as:

$$\text{Acceptor(A)} + \text{Donor(D}^-\text{)} \rightarrow \text{A}^- + \text{D}, \qquad (9)$$

where A and D may be ions or molecules and the reaction takes place in a dielectric solvent. A simple model for electron transfer envisages the electron as being trapped on one of two wells, localized on either A or D, with shapes and relative energies of the two wells determined by the classical degrees of freedom of the nuclei composing A, D and solvent. The electron transfer process between the two wells corresponds to a tunnelling transition. In the limit that the coupling of the electronic (quantum) and classical degrees of freedom is weak, the rate constant for electron transfer can be related to the tunnelling splitting and a Franck–Condon factor. PIMC calculations can be used to calculate the magnitude of the tunnelling splitting[57]. The computation of the tunnelling splitting assumes a large scale separation between the frequency of tunnelling (or inter-well) and intra-well motions. Tunnelling paths correspond to quantum paths which connect the two wells. The fraction of imaginary time spent in the barrier region along such a path will be small because the potential energy in the barrier region is high. This type of rapid process or instanton is then said to correspond to a kink in the quantum path and the tunnelling constant can be related to the free energy required to create a kink in the quantum path. Estimation of this free energy can be accomplished in MC simulations by standard thermodynamic integration methods. The Franck–Condon factor can be related to fluctuations in the relative energy of the two potential wells and estimated from classical MC or MD simulations. An ambitious application of this approach has been in explaining some of the unusual features of the primary electron transfer step in bacterial photosynthesis[58,59]. It should be noted that the above approach estimates a dynamical quantity, the rate constant, in terms of equilibrium thermodynamic averages[60].

This can be shown to be a plausible approximation for many systems and has been applied to calculate rates for quantum processes such as proton transfer and hydrogen chemisorption[61,62].

1. Dirac, P. A. M., *Proc. R. Soc.*, 1929, **123**, 714–735.
2. Franks, F., *Water, A Comprehensive Treatise*, Plenum Press, New York, 1975.
3. Farantos, S. C., Kapetanakis, S. and Vegiri, A., *J. Phys. Chem.*, 1993, **97**, 12158–12166.
4. Parsonage, N. G. and Stavely, L. A. K., *Disorder in Crystals*, Clarendon Press, Oxford, 1978.
5. Allen, M. P. and Tildesley, D. J., *Computer Simulation of Liquids*, Clarendon Press, Oxford, 1987.
6. Putnis, A., *Introduction to Mineral Sciences*, Cambridge University Press, 1992.
7. McCammon, J. A. and Harvey, S. C., *Dynamics of Proteins and Nucleic Acids*, Cambridge University Press, London, 1987.
8. Maitland, G. C., Rigby, M., Smith, E. B. and Wakeham, W. A., *Intermolecular Forces: Their Origin and Determination*, Clarendon Press, Oxford, 1991.
9. Kranenodonk, W. G. T. and Frenkel, D., *Mol. Phys.*, 1991, **72**, 679–713.
10. Payne, M. C., Teter, M. P., Allan, D. C., Arias, T. A. and Joannopoulos, J. D., *Rev. Mod. Phys.*, 1992, **64**, 1045–1097.
11. Ceperley, D. and Alder, B., *Science*, 1986, **231**, 555–560.
12. Schmidt, K. E. and Ceperley, D. M., in *The Monte Carlo Method in Condensed Matter Physics* (ed. Binder, K.), Springer, Berlin, 1992.
13. Hammond, B. L., Lester, W. A. and Reynolds, P. J., *Quantum Monte Carlo Methods in Ab Initio Quantum Chemistry*, World Scientific, Singapore, 1994.
14. Suhm, M. A. and Watts, R. O., *Phys. Rep.*, 1991, **204**, 293–329.
15. Metropolis, N. and Ulam, S., *J. Am. Stat. Assoc.*, 1949, **44**, 335–354.
16. Anderson, J. B., *J. Chem. Phys.*, 1975, **63**, 1499–1053; *J. Chem. Phys.*, 1980, **73**, 3897–4001.
17. Fahy, S., Wang, X. W. and Louie, S. G., *Phys. Rev.*, 1990, **B42**, 3503–3522.
18. Whaley, K. B., *Int. Rev. Phys. Chem.*, 1994, **13**, 41–84.
19. Ceperley, D. M. and Alder, B. J., *Phys. Rev. Lett.*, 1980, **45**, 566–569; Ceperley, D. M., *Phys. Rev.*, 1978, **B18**, 3126–3155.
20. Feynman, R. P., *Statistical Mechanics*, Addison-Wesley, 1972.
21. Doll, J. D., Freeman, D. L. and Beck, T. L., *Adv. Chem. Phys.*, 1990, **78**, 61–127.
22. Ceperley, D. M., *J. Stat. Phys.*, 1991, **63**, 1237–1267.
23. Ceperley, D. M., *Phys. Rev. Lett.*, 1992, **69**, 331–334.
24. Scuseria, G. A., *Nature*, 1993, **366**, 512–514.
25. Marx, D. and Parrinello, M., *Nature*, 1995, **375**, 216–218.
26. Nicholson, D. and Parsonage, N., *Computer Simulation and the Statistical Mechanics of Adsorption*, Academic Press, New York, 1982.
27. Abraham, F. F. and Broughton, J. Q., *Phys. Rev. Lett.*, 1987, **59**, 64–67.
28. Marx, D., Opitz, O., Nielaba, P. and Binder, K., *Phys. Rev. Lett.*, 1993, **70**, 2908–2911.
29. Marx, D., Sengupta, S., Nielaba, P. and Binder, K., *Phys. Rev. Lett.*, 1994, **72**, 262–265.
30. Rossky, P. J. and Schnittker, J., *J. Phys. Chem.*, 1988, **92**, 4277–4285.
31. Hart, E. J. and Boag, J. W., *J. Am. Chem. Soc.*, 1962, **84**, 4090–4095.
32. Chandler, D. and Leung, K., *Annu. Rev. Phys. Chem.*, 1994, **45**, 557–592.
33. Liu, Z. and Berne, B. J., *J. Chem. Phys.*, 1993, **99**, 9054–9064.
34. Fois, E. S., Selloni, A., Parrinello, M. and Car, R., *J. Phys. Chem.*, 1988, **92**, 3268–3273.
35. Thompson, J. C., in *Metal–Non Metal Transitions* (eds Edwards, P. P. and Rao, C. N. R.), Taylor and Francis, London, 1985.
36. Sprik, M., Impey, R. and Klein, M. L., *Phys. Rev. Lett.*, 1986, **56**, 2326–2329.
37. Chandler, D., in *Liquids, Freezing and the Glass Transition* (eds Hansen, J. P., Levesque, D. and Zinn-Justin, J.), Elsevier, Amsterdam, 1991.
38. Deng, Z., Martyna, G. J. and Klein, M. L., *Phys. Rev. Lett.*, 1992, **68**, 2496–2499.
39. Martyna, G. J., Deng, Z. and Klein, M. L., *Chem. Phys.*, 1993, **98**, 555–563.
40. Deng, Z., Martyna, G. J. and Klein, M. L., *J. Chem. Phys.*, 1994, **100**, 7590–7599.
41. Berry, R. S., *J. Phys. Chem.*, 1994, **98**, 6910–6918.
42. Sugano, S., *Microcluster Physics*, Springer, Berlin, 1991.
43. Nayak, S. K., Ramaswamy, R. and Chakravarty, C., *Phys. Rev.*, 1995, **E51**, 3376–3380.
44. Nayak, S. K., Ramaswamy, R. and Chakravarty, C., *Phys. Rev. Lett.*, 1995, **74**, 4181–4184.
45. Goyal, S., Schutt, D. L. and Giacinto Scoles, *J. Chem. Phys.*, 1995, **102**, 2302–2324.
46. Pandharipande, V. R., Zabolitsky, J. G., Piper, S. C., Wiringa, R. B. and Helmbrecht, U., *Phys. Rev. Lett.*, 1983, **50**, 1676–1679.
47. Chakravarty, C., *J. Chem. Phys.*, 1993, **99**, 8038–8043.
48. Chakravarty, C., *J. Chem. Phys.*, 1995, **102**, 956–962.
49. Chakravarty, C., *Mol. Phys.*, 1995, **84**, 845–852.
50. McMahon, M. A., Barnett, R. N. and Whaley, K. B., *J. Chem. Phys.*, 1993, **99**, 8816–8829.
51. Scharf, D., Martyna, G. J. and Klein, M. L., *J. Chem. Phys.*, 1993, **99**, 8997–9012.
52. Scharf, D., Klein, M. L. and Martyna, G. J., *J. Chem. Phys.*, 1992, **97**, 3590–3599.
53. Sindzingre, P., Ceperley, D. M. and Klein, M. L., *Phys. Rev. Lett.*, 1991, **67**, 1871–1874.
54. Scharf, D., Martyna, G. J. and Klein, M. L., *Chem. Phys. Lett.*, 1992, **197**, 231–235.
55. Chakravarty, C., *J. Chem. Phys.*, in press.
56. Chakravarty, C., *Phys. Rev. Lett.*, 1995, **75**, 1727–1730.
57. Marchi, M. and Chandler, D., *J. Chem. Phys.*, 1991, **95**, 889–894.
58. Marchi, M., Gehlen, J. N., Chandler, D. and Newton, M., *J. Am. Chem. Soc.*, 1993, **115**, 4178–4190.
59. Mak, C. H. and Egger, R., *Phys. Rev.*, 1994, **E49**, 1997–2008.
60. Wang, J. and Wolynes, P., *Chem. Phys. Lett.*, 1993, **212**, 427–432.
61. Voth, G. A., Chandler, D. and Miller, W. H., *J. Chem. Phys.*, 1989, **91**, 7749–7760.
62. Mills, G. and Jonsson, H., *Phys. Rev. Lett.*, 1994, **72**, 1124–1127.

Received 18 May 1995; revised accepted 19 November 1995

Vitamin C: A potential saviour against free radical-induced oxidative damage

Indu B. Chatterjee, Chinmay K. Mukhopadhyay and Mrinal K. Ghosh

Dr B. C. Guha Centre for Genetic Engineering and Biotechnology and Department of Biochemistry, University College of Science, Calcutta 700 019, India

Scurvy is the foremost nutritional-deficiency disease that has caused maximum suffering in recorded history. Although research on vitamin C started in 1753, the biochemical functions of the vitamin described so far appear to be nonspecific. This paper is a short review indicating that vitamin C protects mammalian tissues from oxidative damage and that its function is specific. Since oxidative damage is supposed to be a physiological event associated with degenerative diseases and aging, vitamin C appears to be a saviour in protecting us against aging and degenerative diseases.

OXYGEN sustains life, but oxygen is not absolutely friendly. The function of oxygen is to act as a terminal electron acceptor. Usually oxygen receives four electrons on a concerted fashion to produce water, but often it is reduced by single electron to produce oxygen-free radical ($O_2^{\cdot-}$) and reactive oxygen species (ROS), such as H_2O_2, $OH^{\cdot}$ and ferryl. ROS is highly toxic and extremely damaging to biological systems. If not properly scavenged, it results in oxidative damage such as lipid peroxidation, protein oxidation and DNA damage. Oxidative damage has been implicated in several degenerative diseases including inflammation, arthritis, cataract, atherosclerosis, cardiovascular diseases, cancer as well as aging[1-7]. A simple but effective way to prevent these degenerative diseases would be to prevent oxidative damage. A cell's major defence against free radical-mediated damage includes antioxidants like ascorbic acid, vitamin E and glutathione and enzymes like superoxide dismutase, catalase and peroxidases. Sometimes ascorbic acid and vitamin E as well as ascorbic acid and glutathione function as partners in defence[8-11]. Previously, we and others have shown that ascorbic acid is a potential scavenger of $O_2^{\cdot-}$, $^1O_2$, and peroxyl radical[12-15]. In this review, we present a brief account of the discovery of vitamin C, its biosynthesis, and newer insights into its biochemical functions with particular emphasis to indicate that vitamin C (ascorbic acid) is a unique antioxidant and in certain conditions it has a specific role in preventing ROS-mediated oxidative damage of mammalian tissues.

Scurvy and the discovery of vitamin C

Most animals synthesize their own requirement of vitamin C, but a human cannot. Humans are totally dependent upon dietary intake of the vitamin. The knowledge that man cannot produce vitamin C was learnt by experience. But the experience was horrible. During the period of 1500 to 1800 AD, scurvy took a toll of at least two million deaths of sailors[16]. Scurvy also ravaged whole armies and inhabitants of besieged cities. It is perhaps the foremost occupational disease and the nutritional-deficiency disease that has caused most suffering in recorded history[17]. It was the genius of James Lind[18] which banished from the navy the most merciless killer of seamen. In 1753 in his book *A treatise of the scurvy*, Lind pointed out the importance of oranges, lemons and fresh green vegetables in preventing and cure of scurvy. But it took another 175 years to discover the antiscorbutic vitamin by Albert Szent-Györgyi. This discovery was accidental. While extracting and concentrating some redox compound from adrenal glands, Szent-Györgyi isolated some sugar-like crystals about which he was quite ignorant and he called it first 'ignose' [(ig for ignorance and ose for sugar)] and later 'godnose' (God knows)[17]. But the editor of the *Biochemical Journal* objected and he had to give the alternative name hexuronic acid (hex = six). In the same period, Charles Glen King of USA isolated the antiscorbutic factor from lemon juice and named it vitamin C[17]. It showed all the characteristics of Szent-Györgyi's hexuronic acid. It was published in *Science* dated April 1, 1932 entitled 'The chemical nature of vitamin C'. Two weeks later, a note appeared in *Nature* entitled 'Hexuronic acid as the antiscorbutic factor', authored by Svirbely and Szent-Györgyi[17]. But these discoveries were possible only because of the outstanding observation by the Norwegian scientists, A. Holst and his associates. Holst observed that guinea pigs, like humans, were prone to scurvy and could be used for producing experimental scurvy and bioassay of vitamin C. At that time there was quite a flurry as to who should be given the credit for 'discovering vitamin C'. Eventually, Szent-Györgyi received the Nobel Prize for Physiology and Medicine in 1937 in recognition of his discoveries concerning the biological oxidation process with special reference to isolation of vitamin C, but C. G. King, who published five papers and isolated vitamin C, did not. Holst, who made it all possible, died in 1931 without any special honours. Here we see that a scientist, working only for a year in Hopkin's laboratory on a problem with quite

a different objective and without doing a single animal experiment, got the full credit for the discovery of vitamin C. To add to this, in the *Annual Review of Biochemistry* in 1963 (ref. 19), Szent-Györgyi made the comment – 'I was not acquainted with animal tests in this field and the whole problem was for me too glamorous, and vitamins were, to my mind, theoretically uninteresting. Vitamin means that one has to eat it. What one has to eat is the first concern of the chef, not the scientist'.

Biosynthesis of ascorbic acid

While the knowledge of cause and cure of scurvy was clear in 1928–32, the question remained unsolved – How is vitamin C synthesized by most animals and why are humans incapable of producing the vitamin? It took about another thirty years to answer these questions. The most intriguing parts of the problem were the discovery and characterization of two enzymes, namely, aldonolactonase (EC 3.1.1.18, L-gulonolactone hydrolase)[20] and L-gulonolatone oxidase (EC 1.1.3.8, L-gulonolactone : oxygen 2-oxidoreductase) and identification of 2-keto-L-gulonolactone[21–23], an extremely unstable terminal intermediate. Eventually the pathway of biosynthesis of ascorbic acid was completely elucidated by the pioneering work of B. C. Guha (India), C. G. King (USA) and L. W. Mapson (England) and their associates. In animals, ascorbic acid is synthesized from glucose via the glucuronic acid pathway of metabolism[24,25]. In this biosynthetic pathway, L-gulonolactone oxidase is the terminal enzyme. L-gulonolactone oxidase converts L-gulono-1,4-lactone to 2-keto-L-gulonolactone, which is spontaneously converted to ascorbic acid.

Evolution and ascorbic acid biosynthesis

There is a distinct relationship between evolution and the capacity to synthesize ascorbic acid in animals[26–28]. The evolution of L-gulonolactone oxidase (LGO) is species-specific and tissue-specific. LGO is absent in fishes. The enzyme evolved in the kidney of amphibians, resided in the kidney of reptiles, became transferred to the liver of mammals and finally disappeared from the guinea pig, flying mammals, primates including monkeys, apes and humans as well as some highly evolved passeriformes birds[28]. Lack of LGO is the common genetic defect in all these species incapable of synthesizing vitamin C[26,29]. In fact, lack of LGO is an evolutionary loss. Recently, the gene for LGO has been cloned and sequenced by the Japanese group[30–32].

Biochemical functions of vitamin C

Although the pathological syndrome of scurvy was delineated by Lind about 242 years ago, the precise biochemical function of the vitamin remained unclear. One of the reasons is the negative contribution by eminent scientists who had success in one field but were perhaps over-confident in proposing theories in the field of vitamin C which they had not studied in depth in the way they had their original field of work. One single example will suffice to understand this. McCollum, the discoverer of vitamin A, drew the surprising conclusion from his experiments with guinea pigs that it was constipation in the guinea pig, rather than the lack of an unidentified factor in the diet, which was the cause of scurvy[17]. He did not realize that scurvy was a generalized disintegration of all tissues leading to loss of life's total occupation, not to speak only of constipation. In fact, vitamin C had long been neglected. The knowledge that vitamin C is needed for the prevention and cure of scurvy was the cause of ample satisfaction to some people, particularly of the medical profession. Quite a few claims have been made about the biological functions of vitamin C, but most of these are nonspecific[33,34]. Research on vitamin C got a tremendous boost in 1970 and onwards after the publication of a paper in the *Proceedings of National Academy of Science, USA*[35] and a bestseller book, *Vitamin C and the Common Cold*[36] by Linus Pauling. However, Pauling's claim that large doses of vitamin C were beneficial for cold and cancer has not been substantiated by later research. But in spite of the debate, controversy and conflicting results about Pauling's theory, Pauling's message has kept vitamin C research alive. In the five years ending 1981, for instance about 5000 papers were published on vitamin C. But still the precise biochemical function of vitamin C remained controversial. None of the roles of ascorbic acid, described before,[33,34] permits a correlation with the pathologic syndrome found in scurvy, which in reality is a syndrome of generalized tissue disintegration at all levels, both intracellular and extracellular. The generalized tissue disintegration in scurvy is accompanied by subcutaneous and intramuscular haemorrhage in various parts of the body, disruption of collagen being a major cause.

Collagen synthesis

The link between collagen synthesis and vitamin C has dominated our thinking for decades. Collagen is a fibrous connective tissue protein which is abundant in tendons, ligaments, basement membranes, bone matrix and blood vessels. Collagen comprises about one-third of the body protein. The stability of collagen depends upon its triple helix structure, which, in turn depends on the presence of the unique amino acid, hydroxyproline. Ascorbic acid stimulates collagen synthesis by stimulating hydroxylation of proline, which is catalysed by prolyl hydroxylase. Scurvy was thought to be caused due to lack of proper

hydroxylation of proline. However, there is little evidence for direct participation of ascorbic acid in collagen synthesis. The function of ascorbic acid for the stimulation of collagen synthesis is to keep the nonheme iron of prolyl-4-hydroxylase in the active state[37]. During the catalytic reaction, a highly reactive iron–oxygen complex, a ferryl ion, is produced, which subsequently hydroxylates an appropriate proline residue[38]. However, the generation of ferryl ion also proceeds without proline hydroxylation in uncoupled reaction cycles. Ascorbate is utilized as a specific alternative acceptor of the ferryl in these uncoupled reaction cycles. In the absence of ascorbate, prolyl-4-hydroxylase is rapidly inactivated by self-oxidation[38]. It is debatable whether a principal function of ascorbic acid is stimulation of collagen synthesis or prevention of disruption of collagen bundles. Recent evidence from our laboratory indicates that ascorbate protects collagen from ROS-induced oxidative degradation. Moreover, studies both *in vivo* and *in vitro* indicate that ascorbic acid prevents ROS-mediated microsomal lipid peroxidation and protein degradation. These have been discussed later in this paper.

Evolutionary significance

The protective role of ascorbic acid against oxidative damage has evolutionary significance. LGO evolved in the amphibians[26]. The class Amphibia emerged and diversified in the Devonian–Permian, when the atmospheric oxygen concentration increased from 15% to 35% (ref. 39). The water oxygen content was about 0.74% (ref. 39). This would indicate that during evolution of the vertebrates in the terrestrial atmosphere, the tetrapods were exposed to an environmental oxygen concentration of about 47 times that of their aquatic ancestors. This extreme hyperoxia might be a cause of extinction of most of the early amphibians[40]. It is conceivable that only those amphibians that acquired a strong defence mechanism against oxygen toxicity survived successfully. It is known that the form of oxygen responsible for toxicity in hyperoxia is O_2^- and its progeny. The enzymatic defence against O_2^- is superoxide dismutase (SOD)[41–43]. In fact, SOD was contingent for the evolution of early aerobic organisms[41]. In the prokaryotes, SOD is induced in hyperoxia. Exposure of *E. coli* and *S. faecalis* to 25 atm oxygen elicits induction of SOD synthesis 16–25 times, which correlates acquisition of tolerance towards hyperoxia[41]. However, SOD levels of amphibian tissues did not increase over that of fishes[14]. Instead, there was a burst of LGO activity in those successful amphibians[26,27], that survived the hyperoxia through the late Carboniferous and early Permian. The successful tetrapods started synthesizing high amounts of ascorbic acid apparently to protect the tissues against oxidative damage. In the phylogenetic evolution, the biosynthetic capacity continued in the reptiles, most of the mammals and birds.

Subclinical deficiency

Lack of ascorbic acid causes scurvy. So it would appear that if there is no scurvy there is no lack of ascorbic acid. However, scurvy is not a first symptom of lack but a final collapse, a premortal syndrome, and there is a very wide gap between scurvy and full health[44]. Although frank scurvy is rare nowadays, a more incidious condition – the subclinical ascorbic acid deficiency – is common in many parts of the world, particularly involving the nursing mother, infant and elderly. Subclinical deficiency, or marginal ascorbic acid deficiency as it may be called, relates to a state of low tissue level of ascorbic acid, which might cause incipient biochemical changes without showing any apparent clinical symptom. We have recently observed[5,45] that in marginal ascorbic acid deficiency in the guinea pig, oxidative damages such as lipid peroxidation and protein oxidation occur progressively in almost all tissues including liver, lung, heart, kidney, adrenal glands and testes. The oxidative damage occurs despite the presence of adequate levels of other antioxidants like α-tocopherol, glutathione, protein thiols and scavenging enzymes, namely SOD, catalase and glutathione peroxidase[5,45]. Lipid peroxidation and protein changes disappear after ascorbic acid therapy. We have taken the guinea pig as a model animal, because like humans it cannot synthesize ascorbic acid, so the results may be applicable to humans. Particular mention should be made about heart tissue[5]. The levels of antioxidants and free radical scavenging enzymes are normally low in the heart. We have indicated[5] that chronic subclinical ascorbic acid deficiency may produce progressive oxidative damage which in the long run may lead to permanent degenerative vascular diseases and myocardial infarction.

Protection of microsomal membranes against lipid peroxidation and oxidative damage of proteins

Numerous studies[2,46] have been made *in vitro* for elucidating the mechanism of oxidative damage and its prevention. However, most of these studies require the presence of free iron in the incubation medium, a condition which has apparently no relevance to the *in vivo* situation. This is because in the normal physiological condition, iron does not exist in the free state but remains tightly bound with proteins. Using a model *in vitro* system, which is more relevant to the *in vivo* situation, we have demonstrated[47,48] that NADPH initiates lipid peroxidation and oxidative degradation of microsomal proteins in the absence of free iron. Lipid peroxidation and protein

degradation are mediated by cytochrome P450 and specifically prevented by ascorbic acid. Other scavengers of ROS including SOD, catalse, GSH, vitamin E are ineffective. Ascorbate also prevents $O_2^{\cdot-}$-initiated free iron-independent P450-mediated protein oxidation. Probably the oxidant is a perferryl moiety, namely, P450 $Fe^{3+}\cdot O_2^{\cdot-}$. The oxidative degradation of proteins occurs irrespective of tissues, namely liver, kidney, lung, heart, adrenal gland and brain[47,48]. It is a two-step process: (i) $P450Fe^{3+}\cdot O_2^{\cdot-}$-mediated oxidation of proteins and (ii) rapid proteolytic degradation of the oxidized proteins. Using lysyl residue as a model, a mechanism has been proposed[47] for $P450Fe^{3+}\cdot O_2^{\cdot-}$-mediated protein oxidation as indicated below.

$$RCH_2NH_2 + P450Fe^{3+}\cdot O_2^{\cdot-} + H^+ \rightarrow \dot{R}CHNH_2 + P450Fe^{3+} + H_2O_2$$

$$\dot{R}CHNH_2 \rightarrow RCH{=}NH$$

$$RCH{=}NH + H_2O \rightarrow RCHO + NH_3 .$$

Apparently, $P450Fe^{3+}\cdot O_2^{\cdot-}$ abstracts a hydrogen from the side chain amino acid residue producing a carbon-centred radical followed by its conversion to an imino derivative, which then undergoes spontaneous hydrolysis producing an aldehyde derivative of the lysyl residue. Ascorbic acid specifically prevents initiation of protein oxidation by providing an easily donatable hydrogen for abstraction by $P450Fe^{3+}\cdot O_2^{\cdot-}$, thus sparing the side-chain amino acid residue as shown below:

$$P450Fe^{3+}\cdot O_2^{\cdot-} + AH_2 \rightarrow P450Fe^{3+}\cdot O_2^{=} + AH^{\cdot} + H^+$$

$$AH^{\cdot} + AH^{\cdot} \rightarrow AH_2 + A.$$

In this process, ascorbate (AH_2) is oxidized to ascorbate-free radical ($AH^{\cdot}$), which apparently decays by disproportionation to AH_2 and dehydroascorbate (A). Once ascorbic acid intercepts the reaction between $P450Fe^{3+}\cdot O_2^{\cdot-}$ and the side chain amino acid residue, the oxidation of the protein is prevented. So, the subsequent proteolytic degradation of the protein is also prevented.

Protection of collagen against oxidative damage

We have also shown that ascorbic acid prevents oxidative degradation of collagen by ROS produced by $O_2^{\cdot-}$ generated from activated macrophages[49]. The extracellular matrix of mammals contains numerous macrophages that undergo oxidative burst during phagocytosis and release in the environment large amounts of $O_2^{\cdot-}$, redox proteins, like cytochrome b_{558} as well as metalloproteinases. As mentioned above, in the case of microsomal cytochrome P450, collagen is apparently oxidized by a ferryl type of ROS produced from $O_2^{\cdot-}$ and cytochrome b_{558}. The oxidized collagen subsequently undergoes proteolytic degradation by the metalloproteases. Exposure of collagen fibrils to $O_2^{\cdot-}$ generated by the xanthine–xanthine oxidase system also results in rapid oxidation of collagen[49]. In this case, the collagen oxidation is a function of the protein-bound redox iron of xanthine oxidase (XOD). The reactive oxygen species involved is also probably a ferryl moiety, namely, $XOD\text{-}Fe^{3+}\cdot O_2^{\cdot-}$. As observed in the case of microsomal proteins, ascorbic acid apparently prevents collagen oxidation by interacting with the perferryl moiety. Once collagen oxidation is prevented, the subsequent proteolytic degradation of collagen is also prevented. SOD also prevents $O_2^{\cdot-}$-induced collagen breakdown, but in contrast to ascorbic acid which is ubiquitous *in vivo*, the content of SOD in the extracellular fluid is negligible[50,51]. This imparts a specific important role of ascorbic acid for the protection of collagen in the extracellular matrix of mammalian tissues. We consider that an apparently specific function of ascorbic acid is to protect the collagen of the extracellular matrix from oxidative degradation.

Conclusion

The above mentioned results indicate that a potential and apparently specific role of ascorbic acid is to protect the mammalian tissues against oxidative damage both at the intracellular and extracellular levels. We have reported that even in subclinical ascorbic acid deficiency, oxidative damage occurs as evidenced by lipid peroxidation and damage of microsomal proteins[5,45] and that persistent subclinical deficiency results in progressive oxidative damage. We postulate that scurvy symptoms in acute ascorbic acid deficiency is apparently a premortal syndrome of severe oxidative damage, leading to disintegration of tissues. In our present day lifestyle, we are continuously exposed to environmental pollutants such as cigarette smoke, drugs, pesticides, xenobiotics and car exhausts. These pollutants result in increased production of reactive oxygen species, leading to increased oxidative damage such as lipid peroxidation and oxidative degradation of proteins. Since oxidative damage is supposed to be a physiological event associated with various degenerative diseases and aging, we postulate that vitamin C is probably a potential saviour against aging and degenerative diseases.

1. Ames, B. N., *Science*, 1983, **221**, 1256–1264.
2. Halliwell, B. and Gutteridge, J. M. C., *Free Radicals in Biology and Medicine*, Clarendon Press, Oxford, 1989, 2nd edn.
3. Cerutti, P. A., *Science*, 1985, **227**, 375–381.
4. Rath, M. and Pauling, L., *Proc. Natl. Acad. Sci. USA*, 1990, **87**, 6204–6207.
5. Chakraborty, S., Nandi, A., Mukhopadhyay, C. K. and Chatterjee, I. B., *Mol. Cell. Biochem.*, 1992, **111**, 41–47.
6. Mukhopadhyay, M., Mukhopadhyay, C. K. and Chatterjee, I. B., *Mol. Cell. Biochem.*, 1993, **126**, 69–75.

7. Ames, B. N., Shigenaga, M. K. and Hagen, T. M., *Proc. Natl. Acad. Sci. USA*, 1993, **90**, 7915–7922.
8. Chan, A. C., *Can. J. Physiol. Pharmacol.*, 1993, **71**, 725–731.
9. Buettner, G. R., *Arch. Biochem. Biophys.*, 1993, **300**, 535–543.
10. Beyer, R. E., *J. Bioenerg. Biomembr.*, 1994, **26**, 340–358.
11. Meister, A., *J. Biol. Chem.*, 1994, **269**, 9397–9400.
12. Nishikimi, M., *Biochem. Biophys. Res. Commun.*, 1975, **63**, 463–468.
13. Bodannes, R. S. and Chen, P. C., *FEBS Lett.*, 1979, **105**, 195–196.
14. Nandi, A. and Chatterjee, I. B., *J. Biosci.*, 1987, **11**, 435–441.
15. Frei, B., England, L. and Ames, B. N., *Proc. Natl. Acad. Sci. USA*, 1989, **86**, 6377–6381.
16. McCord, C. P., *J. Occup. Med.*, 1959, **1**, 315–318.
17. Carpenter, K. J., *The History of Scurvy and Vitamin C*, Cambridge University Press, New York, 1988.
18. Lind, J., *A Treatise of the Scurvy*, Millar, Edinburgh, 1753.
19. Szent-Györgyi, A., *Annu. Rev. Biochem.*, 1963, **32**, 1–14.
20. Chatterjee, I. B., Chatterjee, G. C., Ghosh, N. C., Ghosh, J. J. and Guha, B. C., *Biochem. J.*, 1960, **74**, 193–203.
21. Chatterjee, I. B., Chatterjee, G. C., Ghosh, N. C., Ghosh, J. J. and Guha, B. C., *Biochem. J.*, 1960, **76**, 279–292.
22. In *Enzyme Nomenclature*, IUPAC and IUB, Elsevier, Amsterdam, 1972, pp. 65, 195.
23. Chatterjee, I. B., Chatterjee, G. C., Ghosh, N. C., Ghosh, J. J. and Guha, B. C., *Naturwissenschaften*, 1959, **46**, 475–476.
24. Chatterjee, I. B., Majumder, A. K., Nandi, B. K. and Subramanian, N., *Ann. N.Y. Acad. Sci.*, 1975, **258**, 24–47.
25. Chatterjee, I. B., *World Rev. Nutr. Diet.*, 1978, **30**, 69–87.
26. Chatterjee, I. B., *Science*, 1973, **183**, 1271–1272.
27. Chatterjee, I. B., *Sci. Cult.*, 1973, **39**, 210–212.
28. Roychoudhuri, C. and Chatterjee, I. B., *Science*, 1969, **164**, 435–436.
29. Chatterjee, I. B., Kar, N. C., Ghosh, N. C. and Guha, B. C., *Nature*, 1961, **192**, 163–164.
30. Khoshizaka, T., Nishikimi, M., Ozawa, T. and Yagi, K., *J. Biol. Chem.*, 1988, **263**, 1619–1621.
31. Nishikimi, M., Koshizaka, T., Ozawa, T. and Yagi, K., *Arch. Biochem. Biophys.*, 1988, **267**, 842–846.
32. Nishikimi, M., Fukuyama, R., Minoshima, S., Shimizu, N. and Yagi, K., *J. Biol. Chem.*, 1994, **269**, 13685–13688.
33. Clemetson, C. A. B., *Vitamin C*, CRC Press, Florida, 1989, vols. I, II and III.
34. Padh, H., *Nutr. Rev.*, 1991, **49**, 65–70.
35. Pauling, L., *Proc. Natl. Acad. Sci. USA*, 1970, **67**, 1643–1648.
36. Pauling, L., *Vitamin C and The Common Cold*, W. H. Freeman and Company, San Francisco, 1970.
37. Kivirikko, K. I. and Myllyla, R., *Methods Enzymol.*, 1987, **144**, 96–114.
38. Tschank, G., Sanders, J., Baringhaus, K. H., Dallacker, F., Kivirikko, K. I. and Gunzler, V., *Biochem. J.*, 1994, **300**, 75–79.
39. Graham, J. B., Dudley, R., Agullar, N. M. and Gans, C., *Nature*, 1995, **375**, 117–120.
40. Erwin, D. H., *Nature*, 1994, **367**, 231–236.
41. Fridovich, I., *Adv. Enzymol.*, 1974, **41**, 35–97.
42. Fridovich, I., *Annu. Rev. Biochem.*, 1975, **44**, 147–159.
43. Fridovich, I., *Annu. Rev. Pharmacol. Toxicol.*, 1983, **23**, 239–257.
44. Szent-Györgyi, A., in *The Healing Factor Vitamin C Against Disease* (ed. Stone, I.), Grosset and Dunlap, New York, 1972, p. xi.
45. Chakraborty, S., Nandi, A., Mukhopadhyay, M., Mukhopadhyay, C. K. and Chatterjee, I. B., *Free Radical Biol. Med.*, 1994, **16**, 417–426.
46. Stadtman, E. R., *Annu. Rev. Biochem.*, 1993, **62**, 797–821.
47. Mukhopadhyay, C. K. and Chatterjee, I. B., *J. Biol. Chem.*, 1994, **269**, 13390–13397.
48. Mukhopadhyay, C. K., Ghosh, M. K. and Chatterjee, I. B., *Mol. Cell. Biochem.*, 1995, **142**, 71–78.
49. Mukhopadhyay, C. K. and Chatterjee, I. B., *J. Biol. Chem.*, 1994, **269**, 30200–30205.
50. McCord, J. M., *Science*, 1974, **185**, 529–531.
51. Marklund, S. L., *Proc. Natl. Acad. Sci. USA*, 1982, **79**, 7634–7638.

ACKNOWLEDGEMENT. We thank the Department of Biotechnology, Government of India, for financial support.

Received 12 June 1995; revised accepted 19 September 1995

RESEARCH ARTICLE

Effect of transient increase in greenhouse gases and sulphate aerosols on monsoon climate

M. Lal*, U. Cubasch, R. Voss and J. Waszkewitz

Deutsches Klimarechenzentrum GmbH, D-20146 Hamburg, Germany
*Permanent address: Centre for Atmospheric Sciences, Indian Institute of Technology, New Delhi 110 016, India

The response to transient increase of greenhouse gases and sulphate aerosols in the Earth's atmosphere on the monsoon climate is examined using the data generated in coupled atmosphere–ocean model experiments performed recently at Deutsches Klimarechenzentrum, Germany. We discuss here the potential role of the sulphate aerosols in obscuring the greenhouse gas-induced warming over the Indian subcontinent during the past century. A comparison of observed and model-simulated trends in annual mean near-surface temperature and monsoon rainfall over the region provides ample evidence of the model's skill. The combined greenhouse gas and sulphate aerosol forcing experiment replicates the observed changes over the region in a broad sense both during winter and on annual mean basis. The year-to-year variability in simulated monsoon rainfall over the Indian subcontinent is in fair agreement with the observed climatology.

We present here a plausible scenario of climate change for the Indian subcontinent as expected by the middle of next century taking into account the projected emission of greenhouse gases and sulphate aerosols. The results suggest an increase in annual mean surface air temperature of 1.0°C over the land regions of the Indian subcontinent in the decade 2040s with respect to the 1980s. During the monsoon season, this warming is found to be less pronounced than in the winter months over the region. Moreover, the warming of land region of the Indian subcontinent is likely to be relatively lower in magnitude as compared to the adjoining ocean, resulting in a decline in land–sea thermal contrast, the primary factor responsible for onset of summer monsoon circulation. As a consequence and contrary to simulations which consider only CO_2 forcing, a weakening of pressure gradient over the region and a decline in monsoon rainfall is simulated by the model.

ANY changes in the radiative balance of the Earth, including those due to anthropogenic increase in greenhouse gases or in aerosols, tend to alter atmospheric temperature and the associated circulation and weather patterns. During the past two decades, several studies have been undertaken to assess the effect of increasing greenhouse gases on the Earth's climate through the use of numerical models[1–8]. The numerical experiments performed earlier in this decade with the state-of-art coupled ocean–atmosphere general circulation models (GCMs) estimated that the increase in greenhouse gases in the atmosphere since the industrial revolution should have generated an increase in global temperature of about 0.6–1.3°C (ref. 9). In reality, the global mean temperatures have risen by about 0.45 ± 0.15°C (ref. 10) which is at the lower end of uncertainty of the response of Earth's climate to greenhouse gases. These experiments predicted a mean rate of global warming of about 0.3°C/decade during the next century as a result of increasing greenhouse gases in the atmosphere[9]. Considerable variability in the rate of warming in different parts of the globe was also suggested.

The most recent IPCC Report[11] has identified sulphate aerosols as a secondary, but potentially important, component that can also affect the climate system by changing the radiative balance of the atmosphere. Sulphate aerosols interact mainly with solar radiation and enhance the Earth's albedo, thereby cooling the surface temperature[12]. In a numerical experiment with an atmospheric GCM coupled to slab ocean using the observed greenhouse forcing over the past 100 years and an estimate of the trend in sulphate aerosol burden calculated by an off-line chemical model, Roeckner *et al.*[13] reported a net cooling over the highly industrialized regions of the northern hemisphere such as eastern United States and large parts of Eurasia. Recent applications of coupled ocean–atmosphere climate models[14,15] have confirmed that, while the combined effect of greenhouse gases and sulphate aerosols is to marginally reduce the mean rate of global warming (0.2°C/decade which is still twice the rate of the tolerance limit of some of the sensitive ecosystems) and that the global warming will accelerate as greenhouse gas forcing begins to dominate the aerosol forcing in coming years, anthropogenic sulphate and biomass combustion aerosols in the troposphere may lead to strong negative forcing in some regions of the globe.

Theoretically, sulphate aerosols are not the only particles that can contribute to cooling. Several kinds of natural aerosols exist in the Earth's atmosphere. They do not, however, seem to be major factors causing change. Natural aerosols – mostly continental dust, sea salt and marine sulphate compounds – have probably

remained constant in their concentration, distribution and properties for at least a century. Thus, they would not have contributed to any observable alterations in climate. The cooling trends precipitated by the major volcanic aerosols, e.g. Krakatoa in 1883 and Pinatubo in 1991 lasted only a few years. In stark contrast, sulphate aerosol compounds in the atmosphere have increased dramatically, primarily during the course of industrialization and most rapidly since about 1950 (refs. 16, 17). Climatologists have thus far focused much of their attention on industrial sulphur compounds as they are the best understood aerosols in view of the large body of data gathered in the past few decades in studies of acid rain. The sulphate aerosols are concentrated over and downwind of the main industrial regions where the present day rate of cooling is found much stronger than the global average and hence locally they may nearly cancel the warming from anthropogenic greenhouse gases. As the anthropogenically induced sulphate aerosol burden in the troposphere would have large spatial and temporal variations in the atmosphere[18], its future impact on regional scales could be in striking contrast to the impact from greenhouse gases for which the concentration changes, in most cases, are likely to be uniform throughout the globe.

Lal *et al.*[19] examined the effect of increasing anthropogenic greenhouse gases (*Business-as-usual* scenario of the Intergovernmental Panel on Climate Change[20]) on monsoon climate using the transient experiments performed with Hamburg coupled atmosphere–ocean (ECHAM1 + LSG) climate model. The study suggested that the mean surface temperature over the Indian subcontinent could significantly rise by $\geq 2°C$ in the next 100 years. An increase in monsoon rainfall over the central plains of India was also suggested. With a view to critically examining the anthropogenic influences on the observed changes in global and hemispheric near-surface temperatures during the past several decades and on the future climate, a set of three new long-term (1880–2049) transient simulations have been performed recently at Deutsches Klimarechenzentrum, Germany[15,21], the first being the control simulation with constant radiative forcings, the second one with observed (for the past century) and projected greenhouse gas-induced radiative forcings and the third with combined observed and projected greenhouse gas and sulphate aerosol forcings. The seasonally and annually averaged spatial patterns of near-surface temperature and precipitation changes are found to be quite different in the experiment with combined greenhouse gas and sulphate aerosol forcings than in the experiment with greenhouse gas forcing alone specially in the northern hemisphere and over Eurasia in particular[21]. In view of this, we have analysed the data generated in these recent climate simulations for a re-assessment of the anticipated climate change over the Indian subcontinent.

In this article, we compare the observed and model-simulated trends in area-averaged annual mean near-surface temperature and monsoon rainfall over the Indian subcontinent for the period 1880–1993. A comparison of observed and model-simulated regional trends in near-surface temperature provides a potentially more stringent test of the model's credibility. An updated plausible future climate scenario for the Indian subcontinent is also presented here.

Climate model

The present study is based on analysis of data generated in numerical experiments performed with the so-called *E*uropean *C*ommunity *HAM*burg (ECHAM version 3) atmospheric model coupled to a *L*arge *S*cale *G*eostrophic ocean model (hereafter referred to as ECHAM3 + LSG model).

ECHAM3 is the third generation atmospheric general circulation model[22] used for global climate modelling investigations in Germany. The model equations are solved on 19 vertical levels in a hybrid sigma-pressure system by using the spectral transform method with triangular truncation at wavenumber 21 (T-21). Nonlinear terms and physical processes, however, are evaluated at grid points of a *Gaussian grid* providing a nominal resolution of 5.625° in latitude and longitude. The radiation scheme is based on a two-stream approximation of the radiative transfer equations with six spectral intervals in the infrared and four in the solar spectrum[23]. Gaseous absorption due to water vapour, carbon dioxide and ozone is taken into account as well as scattering and absorption due to aerosols and clouds. The cloud optical properties are parametrized in terms of the cloud water content. The parametrization of cumulus convection in ECHAM3 version of the model is based on the concept of mass flux scheme of Tiedke[24], which represents various types of convection, from the penetrative convection maintained by large-scale convergence of moisture to the shallow convection such as trade-wind cumuli associated with suppressed conditions. Stratiform clouds are predicted *per se* in accordance with a cloud water equation, including sources and sinks due to condensation/evaporation and precipitation formation. The land surface scheme[25] considers the heat and water budgets in the soil, snow cover and land and the heat budget of permanent land and sea ice. The heat transfer equation is solved in a five-layer model assuming that the heat flux vanishes in the bottom-most layer. The integration is performed following the semi-implicit scheme with leap frog time filter at every 40 min.

The LSG ocean model[26] is based on a numerical formulation of the primitive equations appropriate for large scale geostrophic motion. The nonlinear advection of momentum is neglected and fast gravity waves are

strongly damped by an implicit time integration scheme using a time step of 30 days. The discretization of the ocean model is based on 11 unequally spaced vertical levels and two overlapping 5.625° horizontal grids, which are interpolated onto the resolution of the Gaussian grid used in ECHAM3. In the coupled mode, the basic time step is reduced to 1 day for the computation of sea ice and the temperature and salinity in the two uppermost ocean layers in order to resolve the rapid response of the upper ocean to the short-term variability of the atmosphere.

The ECHAM3 and LSG are coupled by the air–sea fluxes of momentum, sensible and latent heat, short and long wave radiation and fresh water. To avoid a climate drift of the coupled system in long time integrations, a flux adjustment term is applied. Both the models are integrated synchronously but with their respective time steps.

Numerical experiments and region of interest

In the first experiment (hereafter referred to as C), the control reference atmosphere has been simulated with the coupled climate model over a 170-year period with constant atmospheric CO_2 concentration. The second experiment (hereafter referred to as G) included the radiative forcing of observed atmospheric equivalent CO_2 (CO_2 and other greenhouse gases) concentrations for the period 1880–1989 and the projected (based on *Business-as-usual* scenario of IPCC and representing a ~ 1.3% per year compound increase of CO_2) equivalent CO_2 concentrations for the period 1990–2049. In the third experiment (hereafter referred to as S), the effect of both the greenhouse gases (equivalent CO_2) as well as the sulphate aerosols was considered. The anthropogenic sulphate burden information was obtained from Langner and Rodhe's calculations[27] and the direct radiative forcing was mimicked by a change of the surface albedo following the algorithm developed by Charlson *et al.*[12]. The global mean sulphate forcing increased continuously, and more rapidly from 1950 onwards. Figure 1 depicts the spatial distribution of sulphate loading over the Indian subcontinent for the present-day atmosphere and that expected[27] by the middle of next century.

The geographic region of interest for our data analysis reported herein was mainly confined to the region bounded by ~ 5°N to 35°N latitude and ~ 65°E to 95°E (Indian subcontinent and adjoining seas). It may be recalled that the ECHAM models (both version 1 and version 3) have demonstrated[19,28,29] considerable skill in simulating the Indian monsoon circulation and the seasonal cycle of observed climatology for the region at varying resolutions. Therefore, we shall not attempt to elaborate on these general aspects here. It may however be worth mentioning here that the spatial pattern of model-simulated monsoon rainfall in the present version of the model (ECHAM3) is in better agreement

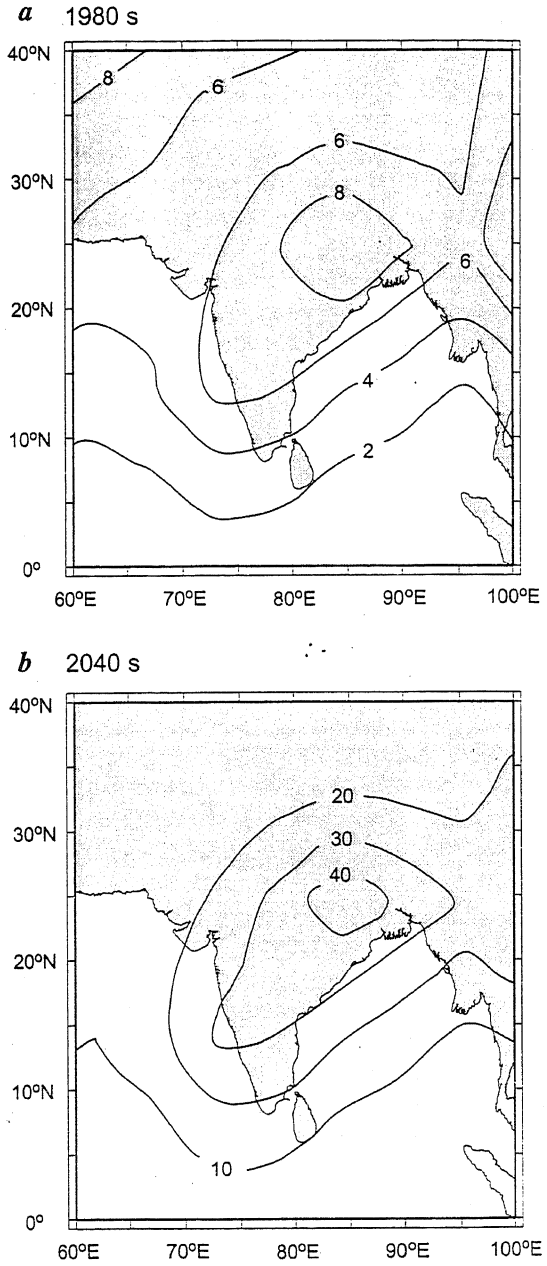


Figure 1. The spatial distribution of anthropogenic sulphate aerosol burden (mg m^{-2}) in the atmosphere over the Indian subcontinent for 1980s and 2040s (based on Langner and Rodhe's calculations[27]).

with the observed pattern (except that the observed sharp east–west gradient in rainfall over peninsular India is not precisely resolved in the model due to rather coarse resolution) than in the previous version of the model (ECHAM1) primarily as a result of inclusion of mass flux convergence scheme.

Comparison of past observed and simulated trends in regional climatology

In order to place a high degree of confidence in future climate change, climate models must replicate the past

trends in key climate elements within the limits of natural variability. Analysis of the recent simulated global and northern hemispheric mean near-surface temperatures has proved[15] that an anthropogenic climate change signal in observed records of near-surface air temperature change can be detected at 95% confidence level. Figure 2 depicts the global 5-year running mean surface air temperatures as observed[30] for the period from 1880 to the present and as simulated by the model for the period 1880–2049 in experiments G and S. An increasing trend in both the observed and the model-simulated annual mean near-surface temperatures is seen in the recent decade but is more rapid in experiment G than either in experiment S or observations. This suggests that model-simulated past trend in global mean surface warming can be reconciled with the observed warming trends only when the cooling effect of aerosols on climate is accounted for.

On regional scales, the higher level of year-to-year variability obscures any systematic differences in surface temperature simulated in experiments G and S in the first half of this century. Figure 3 depicts the 5-year running mean anomalies in area-averaged mean surface air temperatures as observed[31] for the period from 1880 to the present and as simulated by the model for the period 1880–2049 in experiments G and S for the Indian subcontinent (land points only). The temperature anomalies in experiment G show a much higher increasing trend than observed after the year 1970. The temperature anomalies in experiment S do not exhibit any increasing trend until after 1985. The inter-annual variability generated by the model is quite realistic in a statistical sense if we recognize that these climate model simulations do not attempt to precisely simulate individual years.

Table 1 gives the mean annual and seasonal (winter and monsoon) changes in area-averaged near-surface temperature in the decade 1980–89 with respect to the decade 1880–89 over the Indian subcontinent as observed[32] and those simulated in the experiments G and S. It may be noted here that the simulated area-averaged (land only) annual mean temperatures over the Indian subcontinent for the three decades (1931–1960) are 21.7, 21.8 and 21.7°C in experiments C, G and S respectively as compared to the observed[33] climatology of 21.6°C, suggesting that the annual mean near-surface temperature climatology simulated by the model in all the three experiments is quite close to that observed over the region.

Two important points emerge from Table 1. The first is that the annual mean temperature change simulated by the model in the greenhouse gas forcing only experiment is a factor of two higher than the observed changes over the past century over the Indian subcontinent. The combined greenhouse gas and sulphate aerosol forcing experiment is able to replicate the observed changes over the region in a broad sense both during the winter and on annual mean basis. The second point is that the influence of clouds may be more dominant over the region during the monsoon season as evidenced from the observed cooling. An area-averaged change in fractional cloud cover of + 0.3% and – 0.8% is obtained

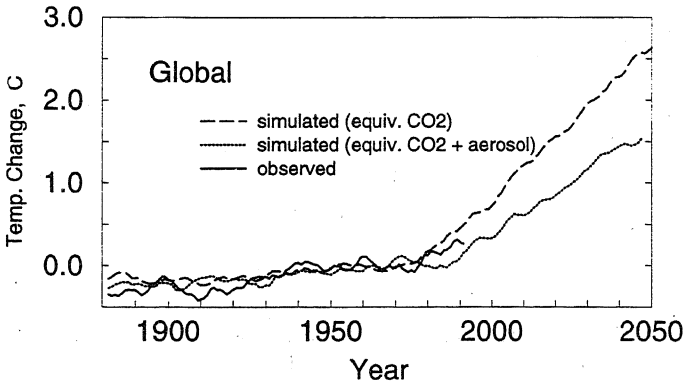


Figure 2. Observed (1880–1992) and model-simulated (1880–2049) trends in global mean surface air temperature anomalies (5-year running means).

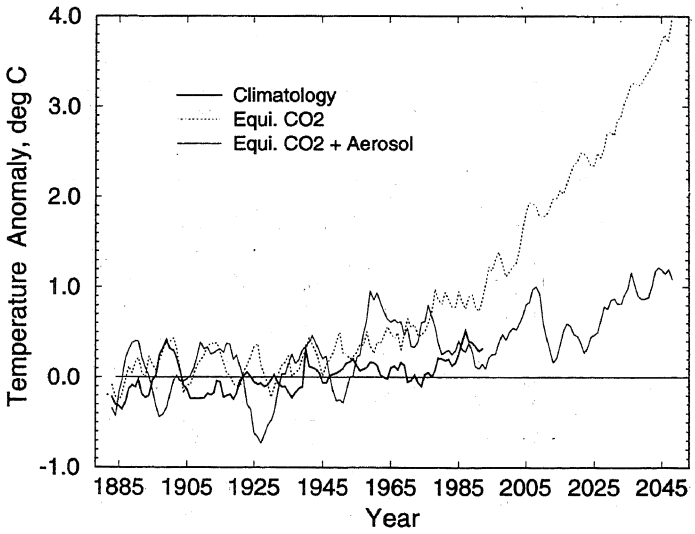


Figure 3. Observed (1880–1993) and model-simulated (1880–2049) trends in area-averaged annual surface air temperature anomalies (5-year running means, land regions only) over the Indian subcontinent.

Table 1. Observed and model-simulated changes in area-averaged near-surface temperature in the decade 1980–89 with respect to the decade 1880–89 over the Indian subcontinent

Period	Experiment G (Land only)	Experiment S (Land only)	Observed
Annual	0.89	0.33	0.4
Winter (DJF)	1.21	0.56	0.7
Monsoon (JJAS)	0.64	0.18	–0.3

in experiments G and S respectively in the decade 1980–89 with respect to the control climate of 1880–89 during the monsoon season over the land regions of the Indian subcontinent. The precise modulating role of clouds in numerical models is not yet fully understood.

The spatial distribution of changes in annual mean near-surface temperature over the Indian subcontinent as simulated in experiments G and S for the decade 1980–89 with respect to the control climate of decade 1880–89 is illustrated in Figure 4. While an increase in annual mean near-surface temperature is simulated everywhere over the Indian subcontinent in experiment G, a decrease in near-surface temperature over the subcontinent to the north of 25°N is simulated in experiment S. While analysing the decadal trends in observed surface air temperature over India, Srivastava *et al.*[34] have clearly demonstrated a decreasing trend in almost all of the northern part of the country (north of 23°N) and a rising trend elsewhere. The simulated cooling over the north and northwest India in experiment S is thus in conformity with the earlier observations[34], indicating that the combined CO_2 and aerosol forcing in the model simulation is able to faithfully reproduce the changes in near-surface temperature over the region as observed during the past century. The simulated warming over the region in experiment G during the past century is much larger than that has occurred in reality. It is interesting to note here that, even though the regionally confined forcing resulting from anthropogenic sulphur emissions has increased relatively sharply only since 1950, its response to the surface temperature change is most pronounced in tropics in a coupled atmosphere–ocean system, in agreement with the current view among scientific community[35]. This suggests that the realized climate change over the region during the past century can, to a large extent, be attributed to the combined greenhouse gas and sulphate emissions.

Figure 5 depicts the 10-year running mean anomalies in area-averaged mean rainfall as observed[36–38] for the period from 1880 to the present and as simulated by the model for the period 1880–2049 in experiments G and S over the Indian subcontinent. The inter-annual

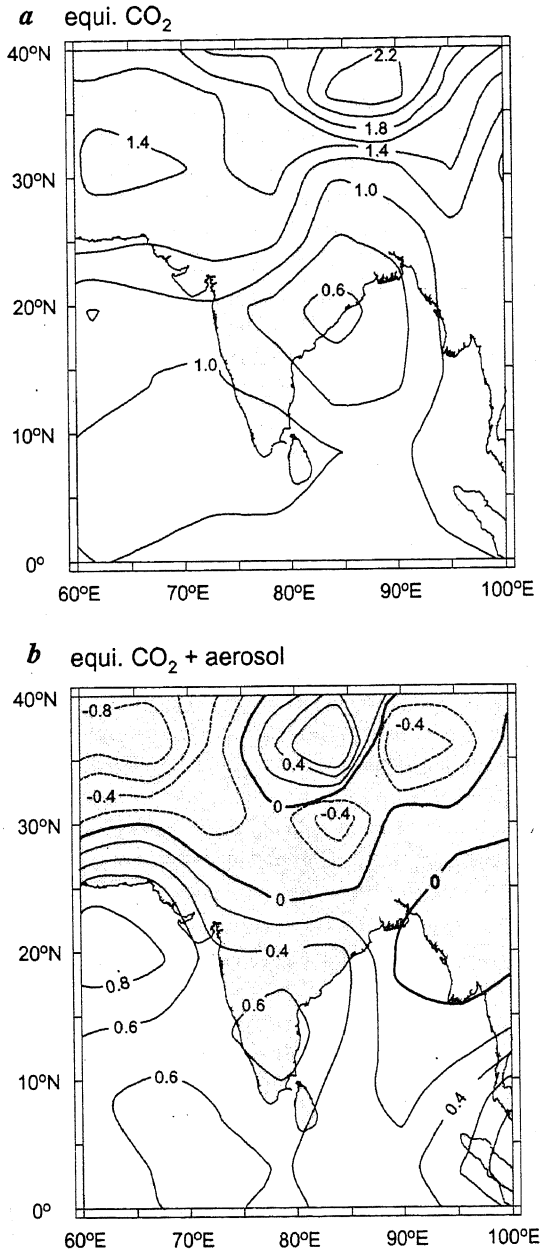


Figure 4. Spatial distribution of change in annual mean surface air temperature (°C) during the decade 1980s with respect to the decade 1880s over the Indian subcontinent as simulated by the model due to (*a*) equivalent CO_2 forcing and (*b*) combined equivalent CO_2 and aerosol forcings.

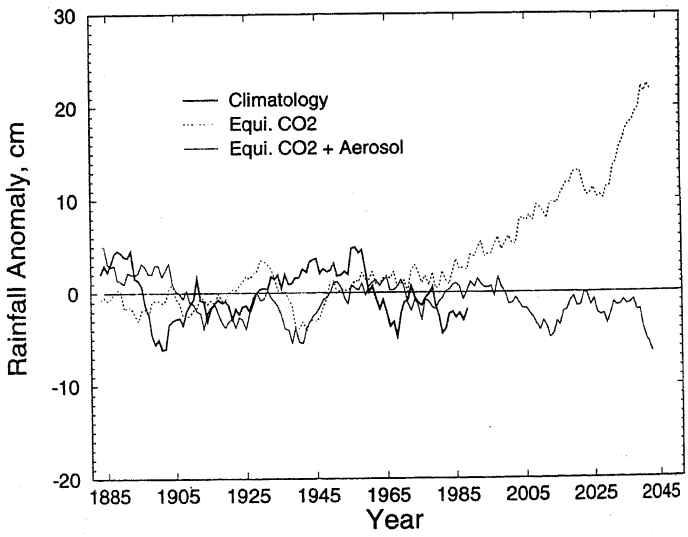


Figure 5. Observed (1880–1993) and model-simulated (1880–2049) trends in area-averaged monsoon rainfall anomalies (10-year running means, land regions only) over the Indian subcontinent.

variability in simulated monsoon rainfall over the Indian subcontinent is in fair agreement with the observed climatology. The simulated rainfall anomalies in the experiment G show an increasing trend after the year 1970. This increasing trend is not discernible in either the observed anomalies or in the experiment S until the present-day.

In the available observed records of rainfall over the Indian subcontinent, a 30-year cyclicity has been reported[32], with no definite trend over the past century. However, as noted in Santer *et al.*[39], while the simulated temperature in a low resolution model on regional scales has a high signal-to-noise ratio, the simulated rainfall is quite noisy and it is difficult to identify any clear increasing/decreasing trends or cyclicity in the simulated rainfall during the past few decades.

The spatial distributions of model-simulated changes in monsoon rainfall over the region in experiments G and S in the decade 1980s with respect to 1880s are depicted in Figure 6. While an increase in excess of 0.5 mm day^{-1} in the past century is simulated in experiment G over the semi-arid regions of north India, both the model simulation in experiment S and the observed records[40] suggest an enhancement in aridity over this region. However, the simulated changes in rainfall over the Indian subcontinent ($\pm$ 1 mm day^{-1}) in both the experiments G and S during the past century are within the range of inter-annual variability observed over the region.

In addition to that discussed above, the ECHAM3 + LSG model has been able to reproduce the observed seasonal reversal in mean sea level pressure, the location of heat low and the other characteristic features of the monsoon circulation over the region considered here in its control simulation. A fair degree of confidence can, therefore, be placed in the projected climate change scenario simulated by the model in experiment S.

Future projections on regional climate change

The time evolution of future changes in globally averaged annual mean surface temperature as simulated in transient experiments G and S is depicted in Figure 2. The model-simulated globally averaged annual mean warming by the decade 2040s with respect to the decade 1980s of 2.48°C in experiment G (with greenhouse gas forcing only) approximates a 0.35°C/decade rate of warming, in overall agreement with the IPCC[20] projection of 0.3°C/decade made for *Business-as-usual* scenario with a simplified upwelling-diffusion model of Wigley and Raper[41]. The globally averaged annual mean warming by the decade 2040s with respect to the decade 1980s simulated by the model in experiment S (combined greenhouse gas and sulphate aerosol forcings) is 1.62°C. This gives a warming rate of 0.23°C/decade over the next century, though lower than with only greenhouse gas forcing but still greater than that observed over the past 10,000 years. The model-simulated rise in globally averaged surface air temperatures during December–February and June–August in the experiments G and S is given in Table 2. It is evident that in both the experiments, the warming is more pronounced during NH winter than in NH summer and over land regions which is typical of all previous modelling results and is physically plausible[9].

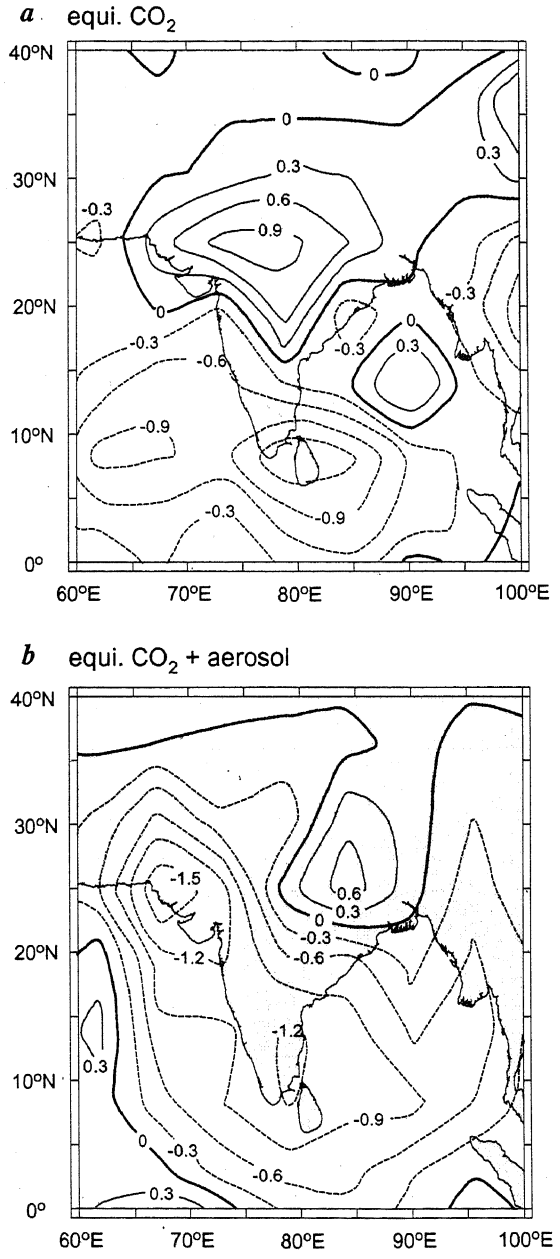


Figure 6. Spatial distribution of change in monsoon rainfall (mm day^{-1}) during the decade 1980s with respect to the decade 1880s over the Indian subcontinent as simulated by the model due to (*a*) equivalent CO_2 forcing and (*b*) combined equivalent CO_2 and aerosol forcings.

The development of useful estimates of future regional scale changes simulated by the global climate model is dependent upon the reliability in simulation of the large-scale circulation features and on the skill of the model in simulating the regional scale past and present-day climatology. The results on the past and present-day climate simulations with ECHAM3 + LSG climate model as discussed above provide reasonable confidence in both these aspects.

The changes in area-averaged annual and seasonal mean 2 m temperature over the Indian subcontinent as simulated by the model in experiments G and S for the period 2040s with respect to 1980s are listed in Table 3. It is seen from this table that while the simulated future warming over the region for experiment G is consistent with other model results[42,43], the warming in experiment S is lower by more than a factor two over the land regions on annual mean basis. Nonetheless, annual mean surface warming over the Indian subcontinent is only marginally lower than that simulated for the globe in spite of increasingly strong negative forcing over the region due to sulphate aerosols.

Another interesting point to note is that while the simulated surface warming is more pronounced over the land region in experiment G (greenhouse forcing only), the land region of the study area is less warm than the adjoining oceanic region in experiment S, more significantly during the monsoon season. This could imply a large-scale change in land–sea thermal contrast that determines relative summer monsoon strength over the region. We shall discuss further on this aspect later.

Table 2. Globally averaged annual and seasonal changes in 2 m temperature as simulated by the ECHAM3 + LSG climate model in G and S experiments

		Equi. CO_2		Equi. CO_2 + aerosol	
Period		2040s–1880s	2040s–1980s	2040s–1880s	2040s–1980s
Land + sea	Annual	2.76	2.48	1.90	1.62
	Winter (DJF)	2.83	2.56	2.00	1.73
	Summer (JJA)	2.72	2.44	1.82	1.54
Land only	Annual	3.61	3.41	2.15	1.95
	Winter (DJF)	3.81	3.55	2.48	2.21
	Summer (JJA)	3.57	3.46	1.93	1.82

Table 3. Model-simulated changes in area-averaged annual and seasonal mean 2 m temperature over the Indian subcontinent for the period 2040s with respect to 1980s in G and S experiments

	Equi. CO_2		Equi. CO_2 + aerosol	
Period	Land + sea	Land	Land + sea	Land
Annual	2.60	2.87	1.41	1.13
Winter (DJF)	2.79	3.16	1.63	1.68
Monsoon (JJAS)	1.87	1.88	1.38	1.05

The spatial distributions of changes in annual mean near-surface temperature over the Indian subcontinent as simulated in experiments G and S for the decade 2040–49 with respect to the present-day climate are illustrated in Figure 7. The simulated surface warming over the region in experiment G during the decade 2040s is ≥ 2°C over most part of the subcontinent, consistent with our findings reported earlier[19]. In experiment S, the magnitude of simulated surface warming

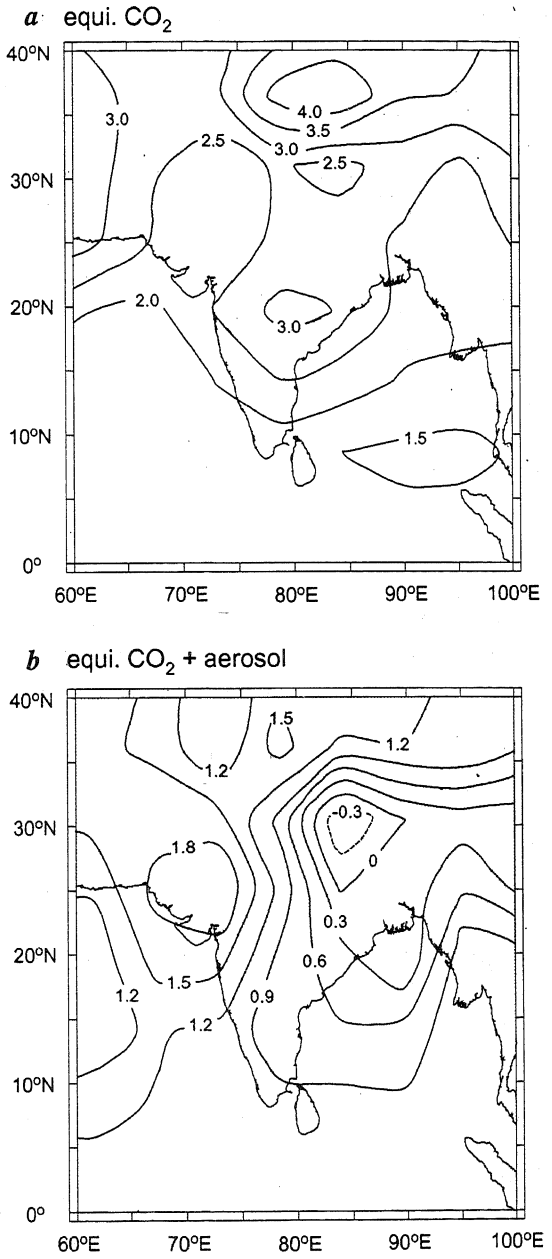


Figure 7. Spatial distribution of change in annual mean surface air temperature (°C) during the decade 2040s with respect to the decade 1980s over the Indian subcontinent as simulated by the model due to (*a*) equivalent CO_2 forcing and (*b*) combined equivalent CO_2 and aerosol forcings.

over most of the region (with the exception of western India) is considerably low (< 1.0°C). A marginal cooling is also simulated in isolated region over the central plains of India.

Figure 8 depicts the spatial distribution of seasonal mean near-surface temperature over the Indian subcontinent as simulated in experiment S for the decade 2040–49 with respect to the present-day climate. During winter, while north and north-west India may experience > 2.0°C warming, cooling of over 2.0°C is simulated by the model over the north-east India. The warming simulated by the model over most of the study area during monsoon season is 1.0°C or less. The surface cooling confined to central India is simulated to be close to 0.5°C during this season.

On global mean basis, the simulated increase in rainfall for the decade 2040s is close to 0.2 mm day^{-1} both during winter and summer in experiment G. This increase is reduced by a factor of two to 0.1 mm day^{-1} during winter and by a factor of four to 0.05 mm day^{-1} during

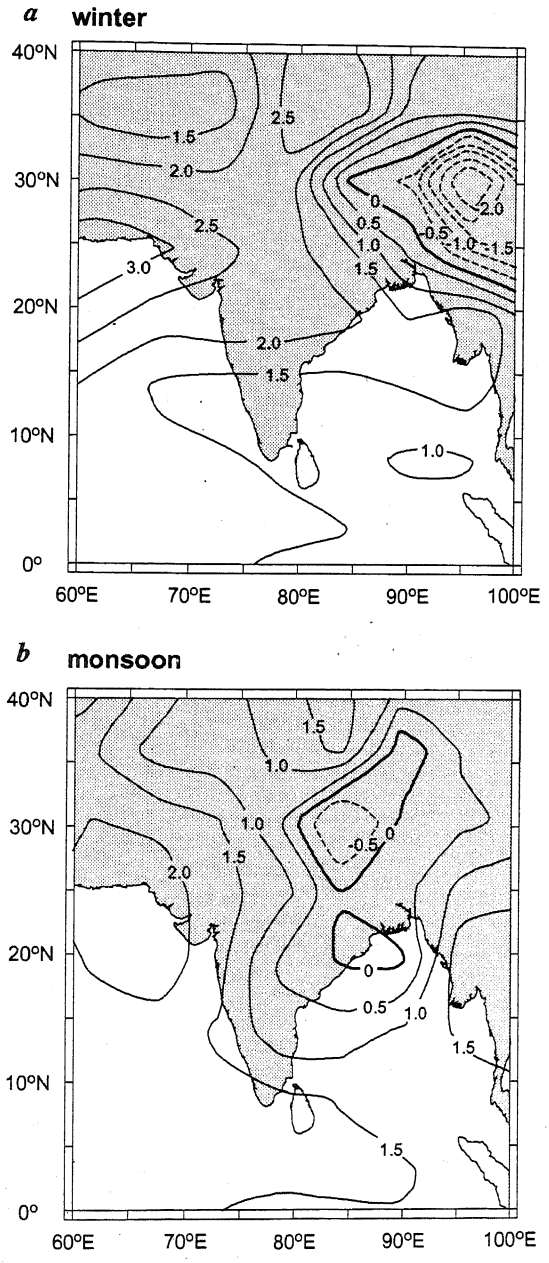


Figure 8. Spatial distribution of model-simulated changes in (*a*) winter and (*b*) monsoon surface air temperature (°C) during the decade 2040s with respect to the decade 1980s over the Indian subcontinent due to combined equivalent CO_2 and aerosol forcings.

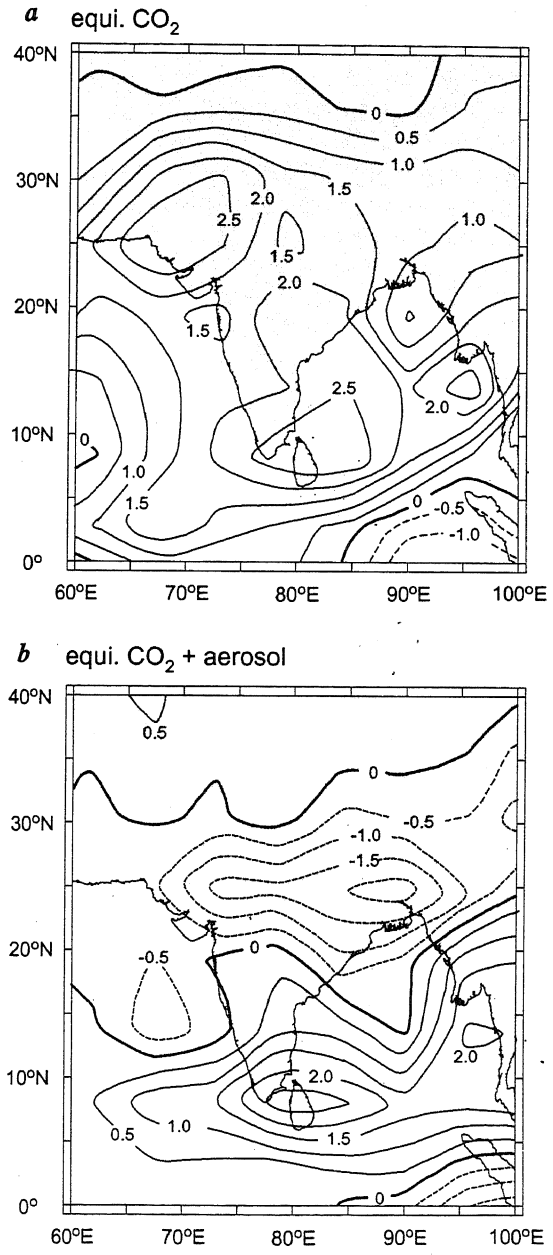


Figure 9. Spatial distribution of change in monsoon rainfall (mm day^{-1}) during the decade 2040s with respect to the decade 1980s over the Indian subcontinent as simulated by the model due to (*a*) equivalent CO_2 forcing and (*b*) combined equivalent CO_2 and aerosol forcings.

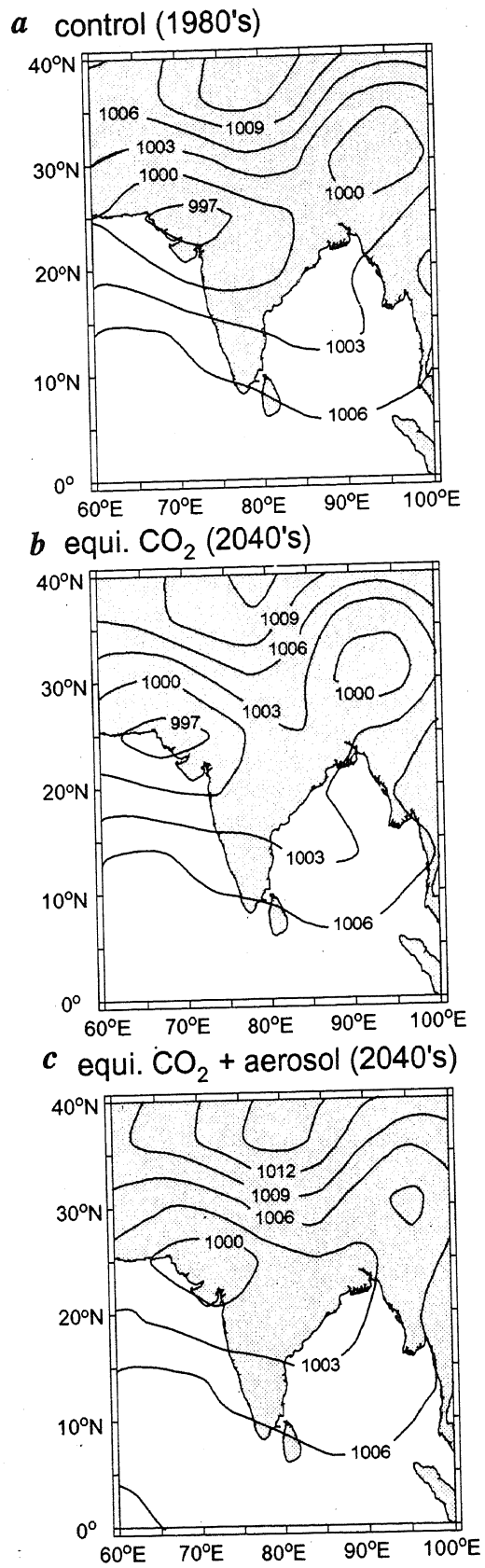


Figure 10. Spatial distribution of model-simulated mean sea level pressure (hPa) over the Indian subcontinent in the month of July during (*a*) 1980s in control experiment, (*b*) 2040s in equivalent CO_2 forcing experiment and (*c*) 2040s in combined equivalent CO_2 and aerosol forcings experiment.

summer in experiment S. While no significant increase in winter precipitation is simulated over the Indian subcontinent, for the decade 2040s, the model simulates an increase in summer monsoon rainfall by about 2 mm day^{-1} in experiment G. The increase in summer monsoon rainfall simulated by the model in experiment G (Figure 5) is consistent with findings from other model simulations[42,43] in enhanced greenhouse gas forcing-only experiments. In experiment S (combined greenhouse gas and aerosol forcings), however, a decrease of about 0.5 mm day^{-1} in summer monsoon rainfall is simulated for the decade 2040s.

The spatial distribution of model-simulated changes in monsoon rainfall over the region in experiments G and S for the decade 2040s with respect to 1980s is depicted in Figure 9. While an increase in excess of 2.5 mm day^{-1} of rainfall is simulated in experiment G over the semi-arid regions of north India with over 1.0 mm day^{-1} elsewhere in the region, the simulated decrease in rainfall over the entire central plains of the Indian subcontinent in experiment S is larger than 1.0 mm day^{-1}. The simulated decline in summer monsoon rainfall over the region in experiment S is found to be marginally above the range of inter-annual variability observed over the region for the present-day atmosphere.

As noted earlier, the simulated increase in surface temperature over the oceanic region of the study area is higher than the land mass of the Indian subcontinent, particularly during the monsoon season in model experiment with combined greenhouse gas and aerosol forcings. This should lead to a reduction in land–sea

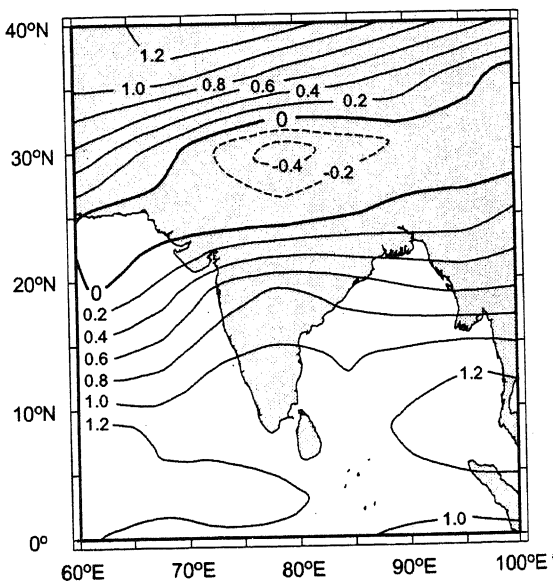


Figure 11. Spatial distribution of change in specific humidity (g/kg) in monsoon season at 850 hPa level during the decade 2040s with respect to the decade 1980s over the Indian subcontinent as simulated by the model in combined equivalent CO_2 and aerosol forcings experiment.

thermal contrast. As a consequence, we expect that the monsoon trough and the heat flow over the north-west India during the monsoon season should weaken. Figure 10 illustrates the spatial distribution of mean sea level pressure in July over the Indian subcontinent for the decades 1980s (in experiment C), and 2040s (in experiments G and S). It is clearly seen that, while the mean sea level pressure distribution over the Indian subcontinent has remained unchanged in experiment G, the low pressure over northwest India has weakened in experiment S. It may be recalled that much of the monsoon rainfall over the central plains of India is associated with the low pressure vortices which develop over the north Bay of Bengal and move on to the subcontinent along the monsoon trough in a northwesterly direction. As a consequence of weaker pressure gradient (and hence an ill-defined monsoon trough), the monsoon circulation simulated in the experiment S is not likely to be as vigorous as simulated in experiment G, leading to lesser moisture convergence over the land regions. A weakening of the low level zonal flow in the decade 2040s relative to decade 1980s is simulated by the model in experiment S over south peninsular India during the monsoon season. The spatial distribution of simulated change in 850 hPa specific humidity during the monsoon season over the Indian subcontinent in the decade 2040s in experiment S (Figure 11) suggests a decline in water vapour content of the atmosphere over the central plains of India. The area averaged simulated change in specific humidity in July over the study region between surface and 500 hPa level is lower by a factor two in experiment S relative to experiment G in the decade 2040s. The relatively less evaporation from the land surface, which is comparatively cooler than in enhanced greenhouse gas forcing only experiment, also reduces water vapour content of the atmosphere, thus resulting in a decline in summer cloud cover (Figure 12) and associated monsoon rainfall over the central plains of India.

An analysis of the area-averaged (land points only) daily rainfall from May to October simulated by the model in experiment C (averaged for the decade 1980s) and experiments G and S (averaged for the decade

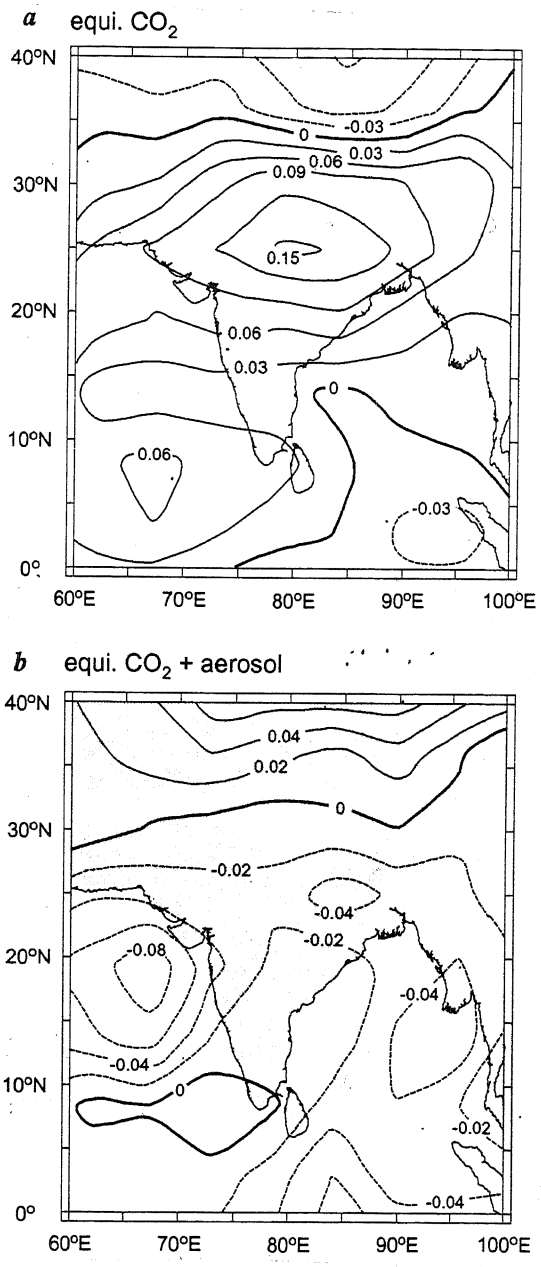


Figure 12. Spatial distribution of change in mean monsoon cloud cover (in fraction) during the decade 2040s with respect to the decade 1980s over the Indian subcontinent as simulated by the model due to (*a*) equivalent CO_2 forcing and (*b*) combined equivalent CO_2 and aerosol forcings.

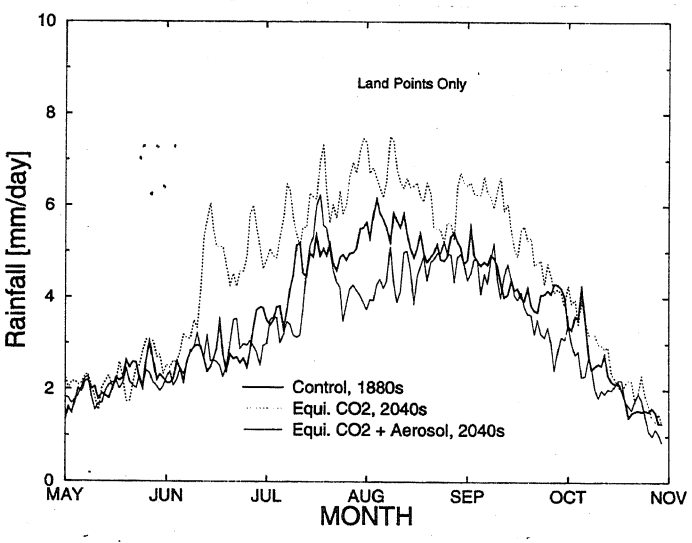


Figure 13. Time evolution of area-averaged (land points only) daily rainfall over the Indian subcontinent during the period May–October in experiment C (1980s) and in experiments G and S (2040s).

2040s) does not suggest early onset or late withdrawal of the monsoon rains over the Indian subcontinent in either of the two anomaly experiments (Figure 13). The simulated decline in monsoon rainfall in experiment S is found to be statistically significant at 90% confidence level. Moreover, while an increased frequency in the intense spells of monsoon rainfall is observed in greenhouse gas forcing experiment, a decline in the frequency of heavy rainfall spells (> 10 mm day^{-1}) is simulated by the model in the combined greenhouse gas and aerosol forcing experiment.

The results presented above highlight the potential influence of combined greenhouse gas and aerosol forcings on the monsoon climate. A more rigorous statistical assessment of the projected changes in monsoon climate and its implications for agriculture and ecosystems will form a separate study.

Conclusions

While there is, as yet, incomplete knowledge of the physical, chemical and radiative properties of aerosols, the available estimates on the mean climatological distribution of sulphate aerosols and the associated 'direct' negative radiative forcings have enabled climate modelling groups in the world (only two till date) to examine its impact on the greenhouse gas-induced warming. The preliminary results have been encouraging in that the scientific community has been able to attribute the observed increases in global mean temperature to anthropogenic activities of mankind with a degree of confidence.

The results obtained in ECHAM3 + LSG model simulation on the past trends in surface air temperature and rainfall are encouraging in that they are able to replicate, in a broad sense, the observed changes in monsoon climate over the Indian subcontinent and this enables us to place a moderate degree of confidence in the simulated climate change scenario presented here. The magnitude of projected surface warming in the decade 2040s over the study region is close to 1.0°C. A decrease of about 0.5 mm day^{-1} in summer monsoon rainfall is simulated by the model in numerical experiment with combined greenhouse gas and aerosol forcings.

It would be interesting to confirm on these findings in other model simulations as and when the required data are made available. A better understanding on the aerosol forcings, increased supercomputing power to perform numerical experiments at finer resolution and improved knowledge of some of the physical processes, which are currently only approximated in parametrization schemes, in coupled atmosphere–ocean climate models are required to provide a more confident regional climate change scenario.

1. Washington, W. M. and Daggupaty, S. M., *Mon. Weather Rev.*, 1975, **103**, 105–114.
2. Manabe, S. and Wetherald, R. T., *J. Atmos. Sci.*, 1975, **32**, 3–15.
3. Manabe, S. and Stouffer, R. J., *J. Geophys. Res.*, 1980, **85**, 5529–5554.
4. Washington, W. M. and Meehl, G. A., *Clim. Dyn.*, 1989, **4**, 1–38.
5. Manabe, S., Spelman, M. J. and Stouffer, F. J., *J. Climate*, 1992, **5**, 105–126.
6. Cubasch, U., Hasselmann, K., Hock, H., Maier-Reimer, E., Mikolajewicz, U., Santer, B. D. and Sausen, R., *Clim. Dyn.*, 1992, **8**, 55–69.
7. Meehl, G. A. and Washington, W. M., *Science*, 1993, **260**, 1101–1104.
8. Murphy, J. M. and Mitchell, J. F. B., *J. Climate*, 1995, **8**, 57–80.
9. Houghton, *et al.* eds, *Intergovernmental Panel on Climate Change 1992*, Cambridge University Press, p. 200.
10. Jones, P. D., Groisman, P. Ya., Coughlan, M., Plummer, N., Wang, W. C. and Karl, T. R., *Nature*, 1990, **347**, 169–172.
11. Houghton, *et al.* eds, *Intergovernmental Panel on Climate Change 1994*, Cambridge University Press, pp. 339.
12. Charlson, R. J., Langner, J., Rodhe, H. and Warren, S. G., *Tellus*, 1991, **43**, 152–163.
13. Roeckner, E., Siebert, T. and Feichter, J., in *Aerosol Forcing of Climate* (eds Charlson, R. J. and Heintzenberg, S.), John Wiley, New York, 1995, pp. 349–362.
14. Mitchell, J. F. B., Johns, T. C., Gregory, J. M. and Tett, S. F. B., *Nature*, 1995 (in press).
15. Hasselmann, K., Bengtsson, L., Cubasch, U., Hergerl, G. C., Rodhe, H., Roeckner, E., Storch, H. V., Voss, R. and Waszkewitz, J., *Geophys. Res. Lett.*, 1995 (communicated).
16. Charlson, R. J., Schwartz, S. E., Hales, J. M., Cess, R. D., Coakley, J. A., Hansen, J. E. and Hoffmann, D. J., *Science*, 1992, **255**, 423–430.
17. Penner, J. E., Charlson, R. J., Hales, J. M., Laulainan, N. S., Leifer, R., Novakov, T., Ogren, J., Radke, L. F., Schwartz, S. E. and Travis, L., *Bull. Am. Met. Soc.*, 1994, **75**, 375–400.
18. Engardt, M. and Rodhe, H., *Geophys. Res. Lett.*, 1993, **20**, 117–120.
19. Lal, M., Cubasch, U. and Santer, B. D., *Curr. Sci.*, 1994, **66**, 430–438.
20. Houghton, *et al.* eds, *Intergovernmental Panel on Climate Change 1990*, Cambridge University Press, pp. 365.
21. Cubasch, U., *Ann. Meteorol.*, 1995, **30** (in press).
22. DKRZ 1994, Technical Report No. 6, DKRZ, Hamburg, pp. 184.
23. Rockel, B., Raschke, E. and Weynes, B., *Bietr. Phys. Atmos.*, 1991, **64**, 1–12.
24. Tiedke, M., *Mon. Weather Rev.*, 1989, **117**, 1779–1800.
25. Dumenil, L. and Todini, E., *EGS Ser. Hydrol. Sci.*, 1992, **1**, 129–157.
26. Maier-Reimer, E., Mikolajewicz, U. and Hasselmann, K., *J. Phys. Oceanogr.*, 1993, **23**, 731–757.
27. Langner, J. and Rodhe, H., *J. Atmos. Chem.*, 1991, **13**, 225–263.
28. Lal, M., *Adv. Atmos. Sci.*, 1994, **11**, 239–246.
29. Lal, M., Bengtsson, L., Cubasch, U. Esch, M. and Schlese, U., *Clim. Res.*, 1995 (in press).
30. Jones, P. D., *Geophys. Res. Lett.*, 1994, **21**, 1149–1152.
31. Rupakumar, K. and Hingane, L. S., *Clim. Change*, 1988, **13**, 521–528 (updated data series from Hingane, L. S. and Ind. Met. Dept., personal communication).
32. Thapliyal, V. and Kulshrestha, S. M., *Mausam*, 1991, **42**, 333–338.
33. Legates, D. R. and Willmott, C. J., *Theor. Appl. Climatol.*, 1990, **41**, 11–21.
34. Srivastava, H. N., Dewan, B. N., Dikshit, S. K., Prakash Rao, G. S., Singh, S. S. and Rao, K. R., *Mausam*, 1992, **43**, 7–20.
35. Stouffer, R. J., Manabe, S. and Vinnikov, K. Ya., *Nature*, 1994, **367**, 634–636.
36. Parthasarathy, B., Rupakumar, K. and Kothawale, D. R., *Meteorol. Mag.*, 1992, **121**, 174–182.
37. Sontakke, N. A., Pant, G. B. and Singh, N., *J. Climate*, 1993, **6**, 1807–1811.

38. Parthasarathy, B., Munot, A. A. and Kothawale, D. R., *Theor. Appl. Climatol.*, 1994, **49**, 217–224.
39. Santer, B. D., Cubasch, U., Mikolajewicz, U. and Hegerl, G., *Clim. Dyn.*, 1994, **9**, 267–285.
40. Chowdhury, A. and Abhyankar, V. P., *Mausam*, 1984, **35**, 375–378.
41. Wigley, T. M. L. and Raper, S. C. B., *Nature*, 1990, **344**, 324–327.
42. Chakraborty, B. and Lal, M., *Terr. Atmos. Ocean. Sci.*, 1994, **5**, 515–536.
43. Bhaskaran, B., Mitchell, J. F. B., Lavery, J. and Lal, M., *J. Climatol.*, 1995, **15**, in press.

ACKNOWLEDGEMENTS. M.L. thanks Deutsches Klimarechenzentrum GmbH (DKRZ) and Max Planck Institute for Meteorology, Hamburg for generous hospitality and for extending facilities to take up this study. The visit of M.L. to DKRZ was funded by the German Ministry for Education, Research and Technology (BMBF). We are grateful to Prof. K. Hasselmann and Prof. L. Bengtsson for their keen interest in this study. Dr E. Roeckner made valuable suggestions during the course of this study. The Meteorological Institute of Stockholm University provided the aerosol forcing data used in the numerical experiments. The assistance of Ms. Marion Grunert in drafting the illustrations is appreciated.

Received 17 May 1995; revised accepted 26 September 1995

RESEARCH COMMUNICATIONS

Surinamite (Be, B, Ga) from cordierite gneisses of eastern ghat mobile belt, India

P. V. Ramesh Kumar, K. K. V. S. Raju and B. S. Ganga Rao*

Department of Geology, Andhra University, Visakhapatnam 530 003, India
*Birla Periclase, Visakhapatnam 530 003, India

Surinamite rich in beryllium, boron and gallium is a monoclinic mineral under group $P2_1/a$ with a = 9.648 Å, b = 11.337 Å, c = 4.981 Å, V = 515.13 Å^3 and Z = 4. Chemical analysis yields the idealized formula as $(Mg, Fe, Mn, Be, Ga)_{3.72}(Al, B)_{3.81}Si_{2.78}O_{15}$. The strongest diffractions in the powder pattern are: (d Å, I, hkl) – 3.7892 (46) (2̄11), 2.6763 (61) (3̄30), 2.436 (100) (041), 2.186 (40) (4̄21), 2.0411 (79) (232), 1.993 (90) (032), 1.8195 (65) (350). Surinamite occurs as an individual greenish blue crystal with characteristic purplish violet pleochroism and has a density of 3.24 g/cc (measured) and 3.27(x) (calculated). It is biaxial negative, $2V_\alpha$ = 68–70°; the refractive indices are α = 1.735, β = 1.739, γ = 1.744; birefringence is moderate to high; $b = Y$; Z = greenish blue, Y = purplish blue and X = brownish violet to yellow. Surinamite (Be, B, Ga) is associated with cordierite, sillimanite, sapphirine, spinel, monazite and quartz. Chemistry indicates couple substitutions as Be and Ga for Mg, Fe, Mn and Al, B for Al, and Al, B for 2 Si. Metamorphic origin has been attributed to this new mineral and such an existence is being reported for the first time from India.

DURING an investigation of sapphirine-associated minerals from cordierite gneisses at various places (locations given in Table 1), we noted some bright greenish blue platy elongated crystals associated with cordierite from eastern ghat mobile belt (EGMB) of Vijayanagaram District, India. About 30 grains of 40 μm size were

Table 1. Chemical analysis, structural formulae of surinamite (Be, B, Ga)

	VN[1] 201	GR[2] 222	AV[3] 235	KK[4] 246	PP[5] 259	PC[6] 263	SN[7] 275
SiO_2	31.89	32.14	32.04	33.15	32.00	32.34	31.94
TiO_2	0.02	0.03	0.02	0.02	0.03	0.02	0.02
Al_2O_3	33.13	33.36	33.16	33.94	33.24	33.54	33.33
Cr_2O_3	0.01	0.01	0.01	0.01	0.01	0.01	0.01
FeO	9.43	9.83	9.75	8.45	9.34	9.15	9.54
MnO	0.07	0.08	0.07	0.08	0.07	0.06	0.06
MgO	16.12	16.54	16.54	16.25	16.34	16.72	16.64
ZnO	0.03	0.03	0.03	0.03	0.03	0.03	0.03
BeO	4.29	4.56	4.49	4.23	4.32	4.61	4.82
B_2O_3	2.56	2.86	2.54	3.17	2.35	2.75	2.01
CaO	0.05	0.05	0.02	0.03	0.04	0.04	0.07
Na_2O	0.01	0.01	0.05	0.04	0.02	0.03	0.02
K_2O	0.01	0.11	0.12	0.14	0.13	0.12	0.17
Ga_2O_3	0.22	0.25	0.27	0.20	0.27	0.23	0.25
Total	97.93	99.86	99.11	99.74	98.19	99.65	98.91
Structural formulae on the basis of 15(O)							
Si	2.79	2.76	2.77	2.82	2.79	2.78	2.77
Ti	0.001	0.002	0.001	0.001	0.002	0.001	0.001
Al	3.41	3.37	3.38	3.40	3.42	3.38	3.41
Cr	0.001	0.001	0.001	0.001	0.001	0.001	0.001
B	0.39	0.42	0.37	0.46	0.53	0.40	0.30
Fe	0.69	0.71	0.71	0.60	0.68	0.65	0.69
Mn	0.005	0.006	0.005	0.006	0.005	0.004	0.004
Mg	2.09	2.11	2.13	2.06	2.12	2.13	2.15
Zn	0.002	0.002	0.002	0.002	0.002	0.002	0.002
Be	0.90	0.94	0.93	0.86	0.91	0.94	1.00
Ga	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Ca	0.005	0.005	0.002	0.002	0.004	0.004	0.006
Na	0.001	0.001	0.008	0.006	0.005	0.005	0.003
K	0.01	0.01	0.01	0.01	0.01	0.01	0.01

[1]Venugopalapuram (lat. 18° 09′ 30″ N, long. 83° 26′ 40″ E)
[2]Gajularega (lat. 18° 07′ 47″ N, long. 83° 24′ 00″ E)
[3]Ayyanna Agraharam (lat. 18° 10′ 44″ N, long. 83° 22′ 43″ E)
[4]Kothuru (lat. 17° 56′ 27″ N, long. 83° 13′ 15″ E)
[5]Panukupeta (lat. 18° 38′ 12″ N, long. 83° 23′ 00″ E)
[6]Padmapuram (lat. 18° 29′ 00″ N, long. 83° 05′ 35″ E)
[7]Sangam (lat. 18° 36′ 12″ N, long. 83° 38′ 50″ E)

separated from cordierite gneiss using Frantz isodynamic separator and finally handpicked using a binocular microscope; the grains were assumed as 99.99% pure. Later subjected to Philips X-ray diffractometer (XRD), the powder pattern resembled that of surinamite as described by de Roever *et al.*[1], suggesting that the mineral was similar to surinamite sensustricto (JCPDS 29-702).

The mineral was found in cordierite gneisses from different localities along with cordierite, sillimanite, green spinel and sapphirine. Accessories comprised ilmenite–magnetite as opaque minerals, monazite and quartz. The sillimanite was present in two generations: (1) rather large platycrystals and (2) small fibrous crystals in aggregates together with the mineral surinamite. Surinamite, in small crystals up to 0.4 mm in size, was found in aggregate together with cordierite and sillimanite. Intergrowths were found between sillimanite and surinamite. About 30–40 grains of surinamite were observed in thin sections of the samples GR 222 and SN 275.

The crystal habit was found to be platy parallel to (010) with one set of well-developed cleavage perpendicular to (010). Optically, surinamite is biaxial negative with $2V_\alpha = 68–70^\circ$ for light 486 μm; refractive indices are: $\gamma = 1.744$, $\beta = 1.739$ and $\alpha = 1.735$. The birefringence is moderate to high. The optical plane is always parallel to (010). The mineral is monoclinic with an optic orientation of $b = Y$. The pleochroic colours of the mineral in almost all the samples are well pronounced as Z = greenish blue, Y = purplish blue expressing dominant purplish glow and X = brownish violet glow to yellow (Figure 1). This is the only mineral expressing glowing brownish violet colour during determination of pleochroism, in contrast to the less pronounced versions of bluish green colours in sapphirine. Hence, the mineral is distinct in its optical characteristics. This variety of the surinamite is reported for the first time from eastern ghat granulite belts of India. The specific gravity was determined to be 3.24 g/cc using the pycnometer technique. The calculated specific gravity from the X-ray data is 3.27(x).

X-ray powder patterns were taken on a Philips X-ray diffraction system PW 1730/PW 1390 with generator setting at 35 kV, 20 mA (Cu-kα radiation having a wavelength of 1.54184 Å; scan speed fixed at 0.04° 2θ/s; 2θ range from 10° to 100°) are presented in Table 2. The X-ray pattern (Figure 2) shows the surinamite to be monoclinic with a possible space group $P2_1/a$. The refined cell parameters[2] were determined to be a_0 = 9.648 Å, b_0 = 11.337 Å, c_0 = 4.981 Å, $\beta = 109° 11'$ and $V = 515.13$ Å^3 with $Z = 4$. The strongest lines are almost all similar to those of surinamite described by de Roever *et al.*[1], except that a few high-angle reflections are less sharp in the pattern for the sample. A few weak reflections are also present in the XRD patterns. Some of the reflections may also belong to the surinamite of the

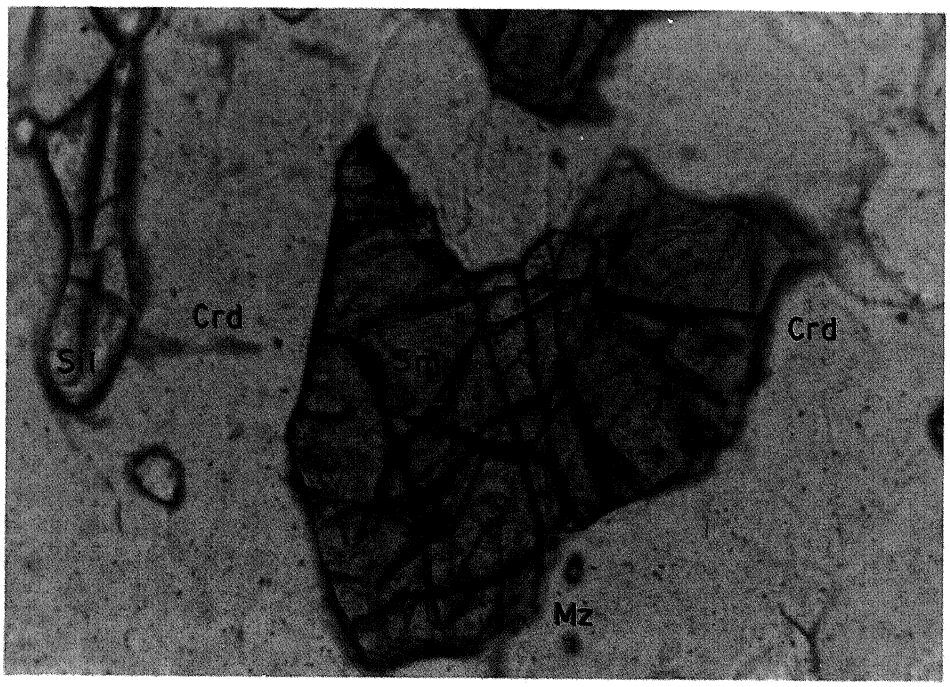


Figure 1. Photomicrograph showing surinamite (Sm) in brownish violet colour to its natural size of 40 μm enclosed within the cordierite (Crd) associated monazite (Mz) crystals with pleochroic haloes along with sillimanite (Sil) (under plane-polarized light 40×).

present study area and could have been obscured in the film of de Roever *et al.*[1] by fluoroscence of Fe excited by Cu radiation.

The crystal structure of surinamite[3] is represented as a polyhedral diagram projected down Z between $0 < Z < \frac{1}{2}$ in Figure 3, based on cubic close-packing of the oxide anions having layers stacked parallel to (001). With essentially Be, B and Ga the cell contents are $M_{24}T_{20}O_{64}$, where M = octahedrally coordinated and T = tetrahedrally coordinated cations by oxygen anions. The atomic occupancy positions for surinamite are given in Table 3. Assuming fully occupied sites, the surinamite cell can have composition $M_9T_5O_{16}$, where M_9 = 3.893 Al + 2.893 Mg + 0.932 Fe + 0.007 Mn + 1.261 Be + 0.04 Ga and T_5 = 3.707 Si + 0.746 Al + 0.547 B is obtained by the additivity rule. The site occupancies of Be and Ga at octahedral positions and B in tetrahedral positions clearly show that no vacancy exists and the structure is fully occupied maintaining the mineral stoichiometry.

Table 2. X-ray diffraction data of surinamite (Be, B, Ga)

d (Å)	I/I_0	hkl	Crystallographic constants
7.081	10	110	System: monoclinic
4.34	24	011	
3.805	15	$2\bar{1}1$	Space group: $P2_1/a$
3.7892	46	$\bar{2}11$	a_0 = 9.468 Å
3.090	20	121	b_0 = 11.337 Å
2.9123	21	230	c_0 = 4.981 Å
2.747	25	$\bar{2}31$	Z = 4
2.6763	61	$\bar{3}30$	D_x = 3.27
2.6417	53	131	
2.436	100	041	Optically negative
2.374	30	330	β = 109° 11′
2.186	40	$\bar{4}21$	V = 515.13 Å^3
2.0411	79	$\bar{2}32$	n_γ = 1.744
1.993	90	032	$2V_\alpha$ = 68–70°
1.9935	75	132	
1.9622	26	$\bar{4}12$	
1.926	10	331	
1.855	18	202	
1.8195	65	350	
1.722	18	$\bar{5}12$	
1.423	16	080	
1.413	18	$\bar{6}32$	
1.391	15	$\bar{6}41$	
1.3445	20	$\bar{2}72$	
1.318	25	352	
1.284	12	—	
1.2256	19	—	
1.222	18	—	

Optically pure separated mineral grains of surinamite were chemically analysed using inductively coupled plasma-mass spectrometer (ICP-MS) (plasmaquad (VG Elemental, Winsford, UK) controlled by an IBM PC-XT computer and associated software) using acceleration potential of RF-1300 W. The carbon-coated samples were checked with the aid of a wavelength-dispersive spectrometer X-ray counting system (WDS XRCS) attached to a scanning electron microscope (SEM) (JEOL JSM-T 330 A) using acceleration potentials of 20 kV. Be, B and Ga were determined from the same sample powders using Perkin Elmer 2380 atomic absorption spectrophotometer (AAS) with acetylene and nitrous oxide as burning fuels. International standards (FeR_2–FeR_4 (G.S. Canada); GIT-IWG of France and G_2 of USGS) were used for analysis with ICP-MS, WDS/SEM and AAS. The results of the complete chemical analysis are presented in Table 1. The average chemical formula calculated on the bases of 15 oxygen atoms is

$$(Mg, Fe, Mn, Be, Ga)_{3.719}(Al, B)_{3.81}Si_{2.78}O_{15} \text{ with } Z = 4.$$

Table 3. Surinamite (Be, B, Ga); atomic occupancy positions

Atom	Occupancy
M(1)	0.517 Mg^{2+} + 0.423 Fe^{2+} + 0.0012 Mn^{2+} + 0.0025 Ga^{2+} + 0.563 Be^{2+}
M(2)	0.692 Al^{2+} + 0.151 Fe^{2+} + 0.0012 Mn^{2+} + 0.0021 Ga^{2+} + 0.1537 Be^{2+}
M(3)	0.768 Al^{2+} + 0.055 Fe^{2+} + 0.0011 Mn^{2+} + 0.0021 Ga^{2+} + 0.1738 Be^{2+}
M(4)	0.811 Al^{2+} + 0.189 Be^{2+}
M(5)	0.841 Mg^{2+} + 0.045 Fe^{2+} + 0.0012 Mn^{2+} + 0.0024 Ga^{2+} + 0.1104 Be^{2+}
M(6)	0.811 Al^{2+} + 0.189 Be^{2+}
M(7)	0.701 Mg^{2+} + 0.207 Fe^{2+} + 0.0012 Mn^{2+} + 0.0024 Ga^{2+} + 0.0884 Be^{2+}
M(8)	0.834 Mg^{2+} + 0.051 Fe^{2+} + 0.0011 Mn^{2+} + 0.0025 Ga^{2+} + 0.1114 Be^{2+}
M(9)	0.811 Al^{2+} + 0.189 Be^{2+}
T(1)	0.155 Al^{3+} + 0.707 Si^{4+} + 0.138 B^{3+}
T(2)	1.000 Si^{4+}
T(3)	1.000 Si^{4+}
T(4)	1.000 Si^{4+}
T(5)	0.591 Al^{3+} + 0.409 B^{3+}
O(1)–O16	Oxygen atoms

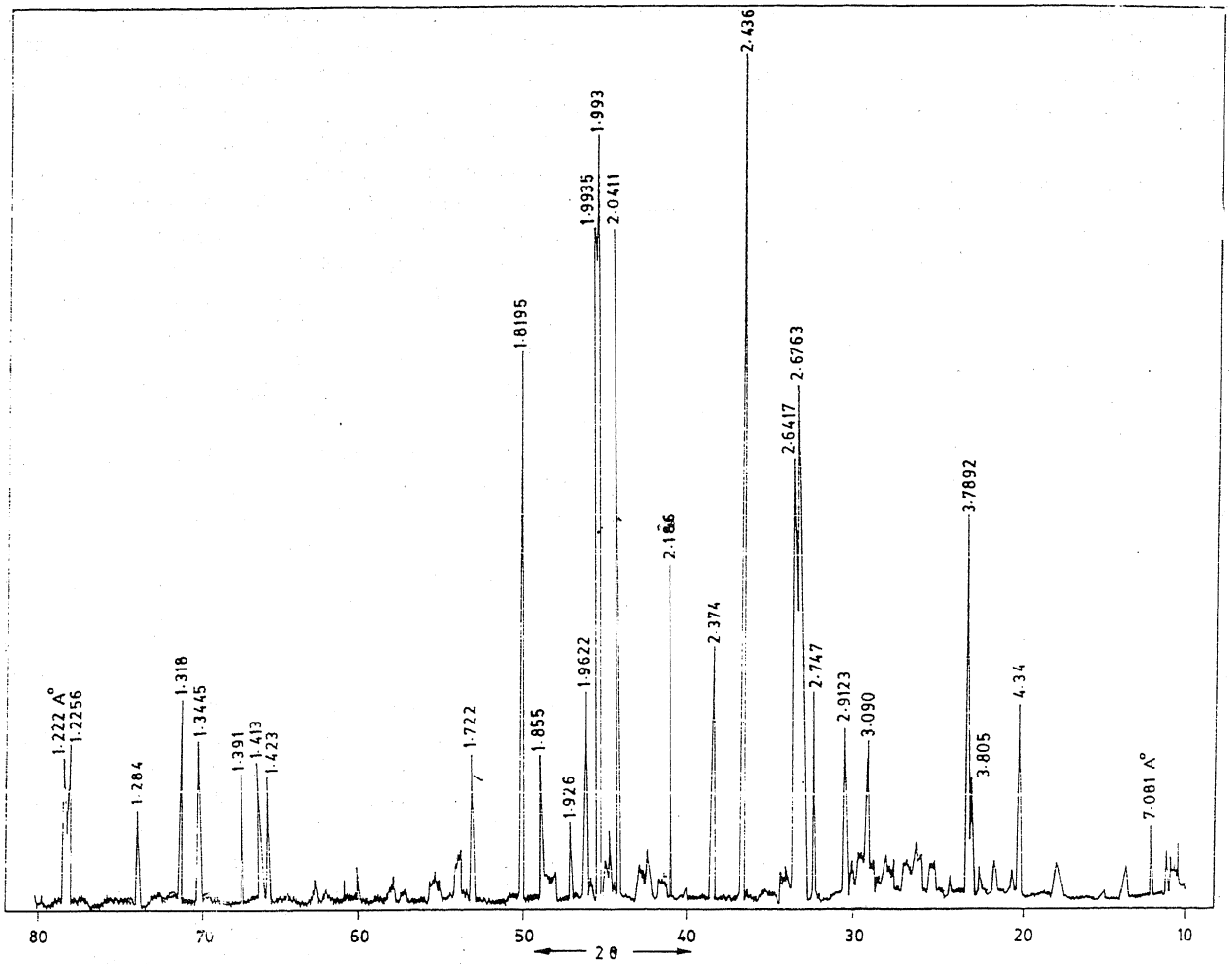


Figure 2. Abridged form of the X-ray diffractogram of surinamite.

The amount brings the analytical totals of the study area samples in the 97.93–99.86% range and the results show reasonable stoichiometry. Clearly, there is a series of cations extending between Mg–Fe–Mn–Be–Ga member and Al–B member of the surinamite. Because Mg, Fe and Mn ions are larger than Be and Ga, and similarly Al and B in the type material. Substitution of a trivalent or bivalent cation for a similar charged cation does not require much compensating mechanism. Exchange operations, if needing coupled substitutions, are technically feasible because the charge compensations can occur in the VIII, VI and IV coordinations for elements like Mg, Fe, Mn, Al which with higher dimensions can easily accommodate Be (0.27 Å) and Ga (0.62 Å) in their respective vacant sites. Charge compensations in the tetrahedral aluminium site could have occurred via the exchange operation with B (0.11 Å). An ionic substitution have both-ways direction (AlSi ↔ SiAl) up to a magnitude to balance the charges and it is logical to treat exchange operators as vectors. In terms of the magnitude, not only can the extent of AlSi four-coordination substitution vary, but it can also be applied to the operation once to change B in the place of substituted Al and Si. Hence, coupled substitutions[4] might have occurred conveniently and compactly as exchange operators, maintaining charge quantities of the elements replaced without disturbing the compact structure of surinamite, and were roughly coeval. Embayed cordierite-associated surinamite and fibrous sillimanite with quartz grain inclusions clearly suggest that breakdown reaction might have occurred at very high pressures during the isobaric cooling of the cordierite gneisses in the area. The reaction may be given as[1]

3 cordierite → 2 surinamite + 2 sillimanite + 7 quartz,

$$3(Mg\ Fe\ Mn)_2\ Al_3\ (Si_5Al)\ O_{18} \rightarrow 2(Mg\ Fe\ Mn)_3\ (Al\ Fe)_4\ Si_3\ O_{15} + 2\ Al_2\ SiO_5 + 7\ SiO_2\ .$$

Some of the chrysoberyl-bearing pegmatites (pegmatite: a plutonic rock of granitic affinity) are the

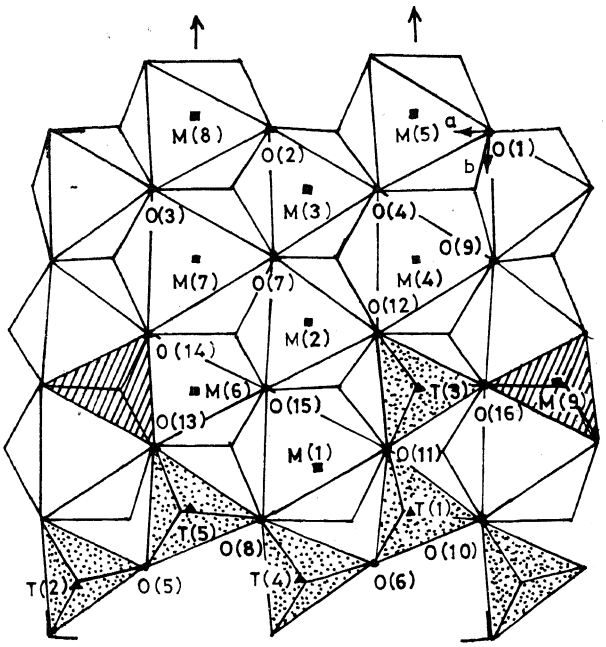


Figure 3. Polyhedral diagram of the surinamite structure. Octahedra in the walls are unshaded except M(9) and the tetrahedra are stippled. The M(9) octahedra between the walls reside above the ruled margin (redrawn after Moore and Araki[3]).

possible source of Be in surinamite at this locality, whereas B and Ga are derived from the original metasediments represented as sillimanite gneisses[5] associated with the khondalites and cordierite gneisses of the study area. Beryllium-, boron- and gallium-bearing surinamite mineralization associated with the cordierite gneisses in a granulite facies terrain may reveal previously unknown geochemical features of the khondalite suite of rocks in the EGMB.

1. de Roever, E. W. F., Kieft, C., Murray, Ed., Klein, E. and Drucker, W. H., *Am. Min.*, 1976, **61**, 193–199.
2. Appleman, D. E. and Evans, H. T. Jr, US Nat. Tech. Inf. Ser. Document PB2-16188, 1970.
3. Moore, P. B. and Araki, T., *Am. Min.*, 1983, **68**, 804–810.
4. Burt, D. M., in *Reviews in Mineralogy* (eds Lipin, B. R. and Mc Kay, G. A.) Mineralogical Society of America, Washington DC, 1989, vol. 21, pp. 259–302.
5. Raju, K. K. V. S., in Proc. Crustal Dynamics, Indian Geophysical Union, Hyderabad, 1987, pp. 107–116.

ACKNOWLEDGEMENTS. We thank the assistance of National Geophysical Research Institute (NGRI), Hyderabad, for ICP-MS analysis; Regional Sophisticated Instrumentation Centre (RSIC), Nagpur, for XRD. We also thank to Dr M. J. Rao, Department of Geology, Andhra University, for assistance and advise during AAS and WDS error-free analyses of the samples.

Received 13 December 1994; revised accepted 28 August 1995

Evidence of garnet to spinel peridotite transition in the harzburgites of Indus ophiolite belt: An indication of their mantle origin

N. R. Karmalkar, A. G. Dessai*, R. A. Duraiswami, M. Leblanc†

Department of Geology, University of Poona, Pune 411 007, India
*Department of Geology, Goa University, Bambolim, Goa 403 202, India
†CNRS, USTL, 34060, Montpellier, France

The transition between garnet and spinel peridotites, caused by the reaction orthopyroxene + spinel = olivine + garnet, has been envisaged in the harzburgites of Indus Ophiolite Belt based on petrographic and geochemical studies. The reaction has been tested in the light of evidence available from phase equilibrium experiments on such transition.

THE harzburgites represent the lowermost section of ultramafic rocks, exposed as tectonically transported materials within the rocks of Dras Volcanic Group[1] or Sangeluma Group[2]. The harzburgite is made up of olivine + orthopyroxene + spinel, with modal abundance of 65 : 32 : 2% and trace amount of clinopyroxene. The harzburgites typically show a transitional texture between protogranular and porphyroclastic texture[3] or porphyroclastic texture[4]. Olivines occur both as porphyroclast (Fo 90.58–92.13) and neoblast (Fo 90.57–91.53). The former varies in grain size from 2 to 4 mm, while the latter varies from 0.5 to 1.5 mm. Majority of the porphyroclasts show development of strain shadows and kink bands. Orthopyroxene porphyroclasts are larger than olivine and vary in size from 3 to 6 mm. Both olivine and pyroxene porphyroclasts have irregular serrated grain boundaries. The pyroxene porphyroclasts are armoured by fine granular aggregates of crushed olivine and pyroxene. The porphyroclasts of orthopyroxene typically show stretching or elongation and are kinked, and often show exsolution lamellae of clinopyroxene. Besides the occurrence of clinopyroxene as exsolution lamellae, it rarely occurs as minute grains within the cluster texture; otherwise they are absent in the rock. Because of their fine grain size their individual microbe analysis is not possible. The contrast in their grain size with the rest of the assemblage and the absence of any deformational features in them raises some doubt as to whether the clinopyroxene is primary. Olivine neoblasts occurring in clusters show a close common orientation, indicating polygonization and recrystallization of former larger grains. One of the significant observations in the harzburgites is that the spinel occurs as fine-grained intergrowths with orthopyroxenes (Figure 1), in a fashion similar to the so-called finger print

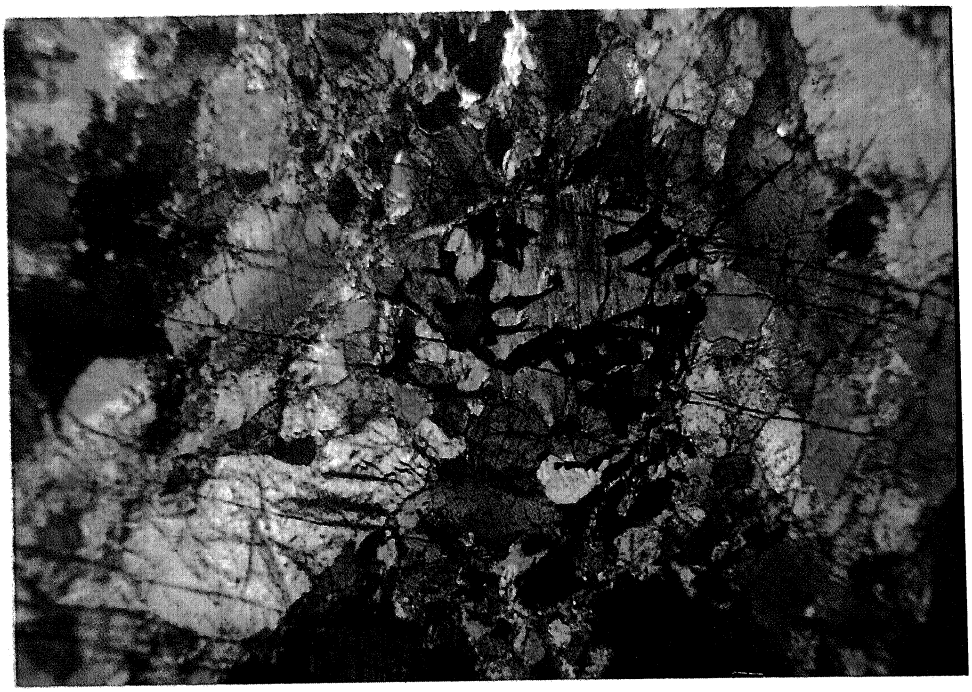


Figure 1. Photomicrograph of harzburgite showing symplectic intergrowth of aluminous spinel with orthopyroxene.

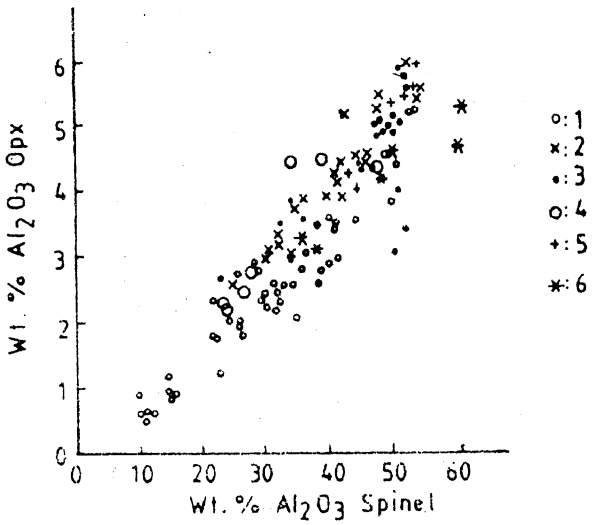


Figure 2. Plot of alumina content of coexisting enstatite and spinel in ophiolitic peridotites. **1**, 'ophiolitic' alpine peridotite; **2**, central Indian ridge; **3**, SW Indian ridge; **4**, mid-Atlantic ridge; **5**, America–Antarctic ridge; **6**, present investigation.

spinels[5]. The proportions of pyroxene and spinel in the cluster cannot be determined accurately because of the irregular boundaries of the aggregates. Though olivine is a dominant phase in the rock it is mostly lacking in the spinel pyroxene graphic intergrowth. The spinel is reddish-brown in colour and besides forming a part of the symplectite intergrowth also occurs as subhedral to euhedral grains interstitial to olivine and orthopyroxene. Concentration of spinel into 'holly leaf' texture was also observed in one of the samples. Composition of orthopyroxenes in clusters is similar to composition of independent orthopyroxene grains in the rock, and any textural explanation should account for the uniformity.

Orthopyroxenes are typically rich in MgO (En 89.21 to 90.32 Fs 7.67 to 9.29 Fm 0.078 to 0.092) content. The Al_2O_3 content of the pyroxene grain ranges from 2.93 to 4.46 wt%. The Mg/Mg + Fe ratio shows a general inverse relationship with the Al_2O_3 content. The spinels are typically rich in Al_2O_3 content, which varies between 32.1 and 53.58 wt% (Table 1). An increase in Cr_2O_3 with decrease in Al_2O_3 was observed form sample no. 32 to 113A, 113B. In the 100 Cr/(Cr + Al) vs 100 Mg/(Mg + Fe) diagram[6] they occupy the field of spinels from ultramafic xenoliths. The Cr and Mg contents behave inversely. The spinel analysis shows them all to be highly aluminous and magnesian, but appreciable amount of chromium and total iron are also present. They are probably best regarded as ferroan chrome spinel.

The strong correlation between Al_2O_3 content of orthopyroxene and spinel found in Figure 2, cannot be just a coincidence but suggests Al partitioning between pyroxene and spinel and supports their co-genetic nature.

The spinel–orthopyroxene symplectite intergrowth texture is seen to be a notable feature of lherzolite/harzburgite nodules as well as alpine peridotites. This texture has been attributed to a variety of origins, viz:

(i) Primary crystallization from pools of melt concentrated during partial fusion[5]: However, pyrometamorphic textures are absent in the studied samples. Besides, the preservation of such original texture seems unlikely considering the deformation that the rock has undergone.

(ii) Mechanical concentration of spinel and pyroxene during deformation[3]: This can explain the concentration of spinel into 'holly leaf' textures. But it cannot explain its strict association only with pyroxene and not with olivine. Alumina partitioned between spinel and enstatite as indicated by the strong correlation cannot be a mere coincidence.

(iii) Exsolution of spinel and diopside from a former aluminous pyroxene[3]: This requires non-stoichiometric conditions for both spinel and pyroxene. Besides, higher values of Al_2O_3 of almost 24 wt% are envisaged for the original pyroxenes. The temperature pressure values required to equilibrate such pyroxenes are unrealistic. The observed high chrome content of the spinel in the sample is in contrast to such an origin.

The absence of plagioclase in the rock under study as well as the low calcic nature of the small clinopyroxene and the high Cr_2O_3 content of the spinel in the cluster texture preclude the possibility of the cluster texture being a product of the reaction between forsterite and anorthite. Similarly, the reaction forsterite + Al-clinopyroxene = spinel + diopside fails to explain the low modal abundance of clinopyroxene in the rock (< 1 vol.%), the strict association of spinel only with orthopyroxene (enstatite) and not with diopside, as well as the observed strong correlation between Al_2O_3 of the orthopyroxene and that of the spinel.

The only feasible explanation is their formation by the reaction

$$\text{pyrope} + \text{olivine} = \text{orthopyroxene} + \text{spinel} + \text{clinopyroxene}^7.$$

Table 1. Electron microprobe microanalysis of orthopyroxenes and aluminous spinels from harzburgites, Ladakh Himalaya

	Orthopyroxene Sample No.					
	113A		113B		32	
	c	m	c	m	c	m
SiO_2	54.66	55.22	56.31	55.78	54.92	55.21
TiO_2	00.03	—	00.06	—	00.10	00.02
Al_2O_3	04.09	03.93	02.93	02.73	04.46	04.16
Cr_2O_3	00.90	00.86	00.66	00.74	00.40	00.55
FeO	05.30	05.35	05.04	05.91	06.10	06.05
MgO	33.64	33.80	34.40	35.53	33.98	34.13
CaO	01.10	01.05	00.76	00.40	00.80	00.47
Na_2O	—	00.02	00.01	00.02	00.02	00.01
K_2O	—	—	00.02	—	—	—
MnO	00.07	00.12	00.13	00.14	00.20	00.12
Total	99.79	100.36	99.32	101.25	100.97	100.72
Wo	02.12	02.01	01.45	00.74	01.50	00.88
En	89.84	89.83	90.82	90.60	89.21	89.98
Fs	08.04	08.16	07.67	08.66	09.29	09.13
Fm	00.0821	00.0833	00.0778	00.0872	00.0922	00.092
t	00.0953	00.0911	00.0675	00.0964	00.0964	00.1049
$t = 0.5\ (Al + Cr - Na)$						
	Aluminous spinels					
TiO_2	—	00.03	00.06	00.06	00.60	00.60
Al_2O_3	44.17	41.71	32.14	34.53	53.58	53.54
Fe_2O_3	01.34	00.46	00.80	01.59	—	—
FeO	13.36	12.54	13.44	13.75	10.23	10.62
Cr_2O_3	24.35	28.15	37.66	35.06	16.75	16.21
MgO	16.56	16.87	15.19	15.56	20.05	19.68
MnO	00.18	00.13	00.23	00.15	00.06	00.10
ZnO	00.23	00.23	00.19	00.07	—	—
Total	100.19	100.12	99.71	100.77	101.27	100.75

This appears to be attendant upon a fall in pressure from garnet peridotite facies to spinel peridotite facies conditions and is consistent with the experimental results[8,9]. The presence of clinopyroxene is additional to the phases in the reaction

pyrope + forsterite = enstatite + spinel

originally proposed[10], due to the addition of CaO to the system[11].

For garnet to react and yield pyroxene with the observed value of $t = 0.06$ the product should have 0.306 molar fraction of spinel[7], which is in good agreement with the measured results (0.30) in the present case. Higher values of t in some samples correspond to a lower modal abundance of spinel in the samples, and larger percentage of independent orthopyroxenes in the rock account for a low modal abundance of spinel. Besides, there is an absence of olivine and garnet in the cluster, indicating the complete reaction of garnet and olivine and conversion of original garnet peridotite assemblage to spinel peridotite. The postulated original garnet may be slightly rich in chrome content, similar in composition to the 'chrome pyrope' (> 4 wt% Cr_2O_3 and higher CaO) variety found in the garnet peridotite xenoliths of Montana diatreams[12]. This will also account for the Cr-rich nature of the spinel and also the small amount of the low calcic clinopyroxene present in the symplectite cluster. The orthopyroxene analyses includes the exsolution lamellaes of clinopyroxene; even then the CaO content in these orthopyroxenes rarely exceeds 1 wt%, which in itself is a proof to consider the low calcic nature of the minute clinopyroxenes in the symplectite. More direct evidences for the olivine–garnet reaction are many[13–15].

The transition from garnet to spinel peridotite as exhibited by the texture may occur during heating or decrease in pressure[13]. Occurrence of pyroxene exsolution lamellae in the rock indicates extensive cooling and so a pressure-induced reaction is more likely[7]. Certainly, such an diapiric movement of the mantle would be in keeping with the evidence of deformation-induced recrystallization textures observed in the sample. It has been shown that the reaction enstatite + spinel = forsterite + pyrope takes place in the pressure range 15–35 kbar at temperatures from 800 to 1600°C (ref. 10). A series of experiments on 'pyrolite' compositions define the stability fields of spinel peridotite and garnet peridotite and the reaction occurs at around 20 kbar in the temperature range 900–1300°C (ref. 16). The boundary is sensitive to the amount of R_2O_3 (Al_2O_3, Cr_2O_3) (ref. 17). Single pyroxene geothermometry[18] gives a temperature range of 1050–1100°C. The pressure values obtained from the Al_2O_3 isopleths and other parameters[19] range between 15 and 25 kbar, which agree well with the experimental values[20,21]. Thus, the garnet–olivine reaction model fits the rock better than any other alternative hypothesis. The reaction described here is reasonably interpreted as an adjustment of the primary high-pressure assemblage to a lower-pressure environment. A decrease in the confining pressure from 50 to below 20 kbar would move the garnet out of its stability field[14].

Garnet peridotite would be subject to pressure decrease during tectonic uplift or mantle upwelling; this may have been responsible for the formation of such symplectite texture. The cluster texture may, therefore, be an indicator of mantle convection. Occurrence of spinel–pyroxene symplectite intergrowth in the alpine peridotite or ultramafic xenolith can be taken as an evidence of garnet to spinel peridotite transition in the mantle.

1. Searle, M. P., *Trans. R. Soc. Edin. Earth. Sci.*, 1983, **73**, 205–219.
2. Srikantia, S. V. and Razdan, M. L., *J. Geol. Soc. India*, 1980, **21**, 523–545.
3. Mercier, J. C. C. and Nicolas, A., *J. Petrol.*, 1975, **16**, 454–487.
4. Pike, J. E. and Schwarzman, E. C., *J. Geol.*, 1976, **85**, 49–61.
5. Dawson, J. B. and Smith, J. V., *Phys. Chem. Earth*, 1975, **9**, 339–350
6. Irvine, T. N. and Findley, T. C., in *The Ancient Ocean Lithosphere*, Dept. E.M.S., 1972, vol. 43, pp. 97–126.
7. Smith, D., *J. Geol.*, 1977, **85**, 476–482.
8. MacGregor, I. D., *J. Geol.*, 1965, **64**, 126–134.
9. Cohen, L. H., Ito, K. and Kennedy, G. C., *Am. J. Sci.*, 1967, **265**, 475–518.
10. MacGregor, I. D., *Yb. Carnegie Instn. Washington*, 1964, **63**, 157.
11. Dawson, J. B., Powel, D. G. and Reid, A. M., *J. Petrol.*, 1970, **11**, 519–548.
12. Heran, B. C. Jr, and McGee, E. S., *US Geol. Surv. Bull.*, spl. paper, 1983, **1604**, 1–33.
13. Wallace, R. C., *Contrib. Mineral. Petrol.*, 1975, **49**, 191–199.
14. Reid, A. M. and Dawson, J. B., *Lithos*, 1972, **5**, 115–124.
15. Varne, R., *J. Petrol.*, 1977, **81**(1), 1–23.
16. Green, D. H. and Ringwood, A. E., *Earth Planet. Sci. Lett.*, 1967, **3**, 151–160.
17. MacGregor, I. D., *Phys. Earth Planet. Inter.*, 1970, **3**, 372–377.
18. Mercier, J. C. C., *Am. Mineral.*, 1976, **61**, 603–661.
19. Mysen, O. B., *Am. Mineral.*, 1976, **61**, 677–683.
20. MacGregor, I. D., *Am. Mineral.*, 1974, **59**, 110–119.
21. Dankwerth, P. A. and Newton, R. C., *Contrib. Mineral. Petrol.*, 1978, **66**, 189–201.

ACKNOWLEDGEMENT. We thank CSIR, New Delhi, for financial assistance.

Received 22 December 1994; revised accepted 11 September 1995

Bilateral outgrowths on thorax of sepia eye colour mutants in *Drosophila bipectinata*

B. N. Singh, Seema Sisodia and Rakhee Banerjee

Department of Zoology, Banaras Hindu University, Varanasi 221 005, India

Flies with bilateral outgrowths on thorax were detected in sepia eye mutant stock of *Drosophila bipectinata*. A separate stock of sepia mutants with outgrowths on thorax could be established. Expression of this phenotypic change is better in females than in males and there is individual variation in both sexes. In order to understand the mode of inheritance of these outgrowths on the thorax of sepia mutants, four different crosses involving mutants and wild type flies were made. The proportion of flies with outgrowths on thorax varies in F_1 and F_2 generations of different crosses. Such outgrowths on the thorax have not been reported earlier in any species of *Drosophila* and appear to be a unique phenotypic change in *D. bipectinata* caused due to mutation but its exact nature could not be understood from the results of these crosses.

D. BIPECTINATA is a member of the *bipectinata* species complex of the *ananassae* subgroup of the *melanogaster* species group. This species is of common occurrence in India. Studies on population and behaviour genetics of this species have been initiated by Indian workers[1-9]. We detected sepia eye colour mutation (autosomal reces-

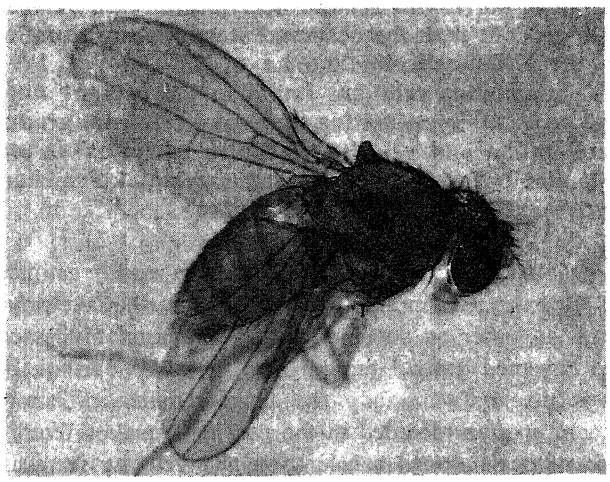


Figure 1. Photograph showing sepia eye mutant of *D. bipectinata* with bilateral outgrowths on thorax.

Table 1. Results of crosses involving sepia eye colour mutants with/without outgrowths on thorax and wild type *Drosophila bipectinata*

1.	sepia eye colour ♀ normal thorax			×	sepia eye colour ♂ thorax with outgrowths			
	F_1							
	sepia eye colour normal thorax				sepia eye colour thorax with outgrowths			
	♀	♂	Total		♀	♂	Total	
	140	159	299		94	99	193	
	F_2							
	sepia eye colour normal thorax				sepia eye colour thorax with outgrowths			
	♀	♂	Total		♀	♂	Total	
	38	50	88		74	60	134	
2.	sepia eye colour ♀ thorax with outgrowths			×	sepia eye colour ♂ normal thorax			
	F_1							
	sepia eye colour thorax with outgrowths				sepia eye colour normal thorax			
	♀	♂	Total		♀	♂	Total	
	106	53	159		80	72	152	
	F_2							
	sepia eye colour thorax with outgrowths				sepia eye colour normal thorax			
	♀	♂	Total		♀	♂	Total	
	98	86	184		98	124	222	
3.	Red eye colour ♀ normal thorax			×	sepia eye colour ♂ thorax with outgrowths			
	F_1							
	Red eye colour normal thorax				Red eye colour thorax with outgrowths			
	♀	♂	Total		♀	♂	Total	
	157	204	361		0	0	0	
	F_2							
	Red eye colour normal thorax				Red eye colour thorax with outgrowths			
	♀	♂	Total		♀	♂	Total	
	91	71	162		8	11	19	
	sepia eye colour normal thorax				sepia eye colour thorax with outgrowths			
	♀	♂	Total		♀	♂	Total	
	10	16	26		18	21	39	
4	sepia eye colour ♀ thorax with outgrowths			×	Red eye colour ♂ normal thorax			
	F_1							
	Red eye colour normal thorax				Red eye colour thorax with outgrowths			
	♀	♂	Total		♀	♂	Total	
	218	187	405		7	0	7	
	F_2							
	Red eye colour normal thorax				Red eye colour thorax with outgrowths			
	♀	♂	Total		♀	♂	Total	
	157	143	300		28	15	43	
	sepia eye colour normal thorax				sepia eye colour thorax with outgrowths			
	♀	♂	Total		♀	♂	Total	
	19	32	51		28	10	38	

sive) in a wild stock of *D. bipectinata*[10] which diminishes sexual activity of males[8]. In sepia mutant stock, 17 females and 10 males showing two outgrowths on dorsal side of thorax were found. Figure 1 shows the photograph of a sepia mutant with two outgrowths on the dorsal side of the thorax. A separate line of sepia mutants showing outgrowths on thorax could be established. It has been observed that there is individual variation in both sexes in expression of this character and generally outgrowths are larger in females than in males. In the stock of mutants showing outgrowths, a few flies are regularly found which do not possess outgrowths on their thorax. In the sepia mutant stock also, a few flies with outgrowths are regularly found.

To understand the mode of inheritance of this unique phenotypic change, four different crosses were made. All the crosses were made in food bottles by using wild type, sepia mutants with and without outgrowths on thorax (normal thorax). In each cross, a large number of F_1 flies were observed to score the number of flies with/without outgrowths on thorax. A random sample of 50 flies (females and males in equal number) from F_1 were transferred to a fresh food bottle to obtain the F_2 generation of each cross. F_2 flies were observed to score the different types of flies. Results of all four crosses are presented in Table 1. In reciprocal crosses (1 and 2) between sepia mutants with normal thorax and with outgrowths, both types of flies were found in F_1 and F_2 generations and the results of reciprocal crosses are more or less similar. The frequency of sepia mutants with outgrowths on thorax varies from 40 to 60% in F_1 and F_2 of these crosses. However, the results of reciprocal crosses between sepia mutants with outgrowths and wild type (red eye and normal thorax) flies are different from those of the crosses involving sepia mutants with normal thorax and with outgrowths on thorax. When wild type females were crossed with sepia mutant males with outgrowths, flies with outgrowths on thorax were absent in F_1. In the F_2 generation of this cross, red-eyed flies with outgrowths and without outgrowths as well as sepia mutants with outgrowths and without outgrowths were found but the frequency of flies with outgrowths was lower (23.58%) as compared to the crosses involving sepia mutants. In the opposite cross (sepia females with outgrowths × wild type males), seven females with very small outgrowths on thorax were found in the F_1 generation. In the F_2 generation, four types of flies were found and the frequency of flies with outgrowths on thorax was very low (16.75%). Thus the results of different crosses vary and development of outgrowths on thorax occurs in both wild type (red eye) and sepia mutants of *D. bipectinata*.

1. Singh, B. N. and Das, A., *Biol. Zent. bl.*, 1991, **110**, 157–162.
2. Singh, B. N. and Chatterjee, S., *Evol. Biol.*, 1991, **5**, 105–113.
3. Singh, B. N. and Chatterjee, S., *Genome*, 1991, **34**, 849–852.
4. Singh, B. N. and Pandey, M., *Indian J. Exp. Biol.*, 1991, **29**, 926–929.
5. Naseerulla, M. K. and Hegde, S. N., *Indian J. Exp. Biol.*, 1993, **31**, 215–218.
6. Singh, B. N. and Sisodia, S., *Biol. Zent. bl.*, 1995, **114**, 95–101.
7. Banerjee, R. and Singh, B. N., *Proc. Zool. Soc. Calcutta*, 1995, **47**, 125–133.
8. Sisodia, S. and Singh, B. N., Third *Drosophila* Meeting, Mysore, Abstract, 1995, p. 37.
9. Sisodia, S. and Singh, B. N., *Zoological Studies*, 1995, **35**, in press.
10. Singh, B. N., Sisodia, S. and Banerjee, R., *Dros. Inf. Serv.*, 1995, in press.

ACKNOWLEDGEMENTS. Financial support from the UGC, New Delhi, in the form of research project to BNS, a JRF to RB and from BHU in the form of research scholarship to SS is gratefully acknowledged. We also thank Dr A. K. Sundaran for his help in photography.

Received 19 July 1995; accepted 29 September 1995

A new empirical relation for strong seismic ground motion for the Himalayan region

Ashutosh Aman, Umesh K. Singh and Ramesh P. Singh

Department of Civil Engineering, Indian Institute of Technology, Kanpur 208 016, India

Strong ground motion data from five earthquakes that occurred along the Himalayan range have been taken to study the attenuation characteristics of horizontal peak acceleration and velocity. Using these data, we have found empirical relations for horizontal acceleration and velocity. These empirical relations for peak horizontal ground acceleration and velocity give good fit with the observed strong ground motion in the Uttarkashi earthquake of 20 October 1991. The attenuation relations found in the present study will be useful in estimating the peak horizontal ground acceleration and velocity of future earthquakes and will also be useful in designing earthquake-resistant structures in the Himalayan region.

EARTHQUAKE-resistant design of structures require estimate of the expected ground motions at the earthquake-prone sites. In order to develop ability to reasonably assess the expected ground motion due to future earthquake, it is essential to study the characteristics of strong ground motion recorded in the past earthquakes. Prediction of peak ground acceleration and velocity is useful to obtain an estimate of the outer envelope of the ground motion spectrum[1]. Peak ground motions are important in earthquake safety analysis and in the evalua-

Table 1. Earthquakes considered in the present work

Earthquake	Date of occurrence	Magnitude (M_B)
Dharamsala	26 April 1986	5.7
Meghalaya	10 September 1986	5.7
Burma–India	14 May 1987	5.7
Tripura–Assam	6 February 1988	5.8
Gauhati	6 August 1988	7.2

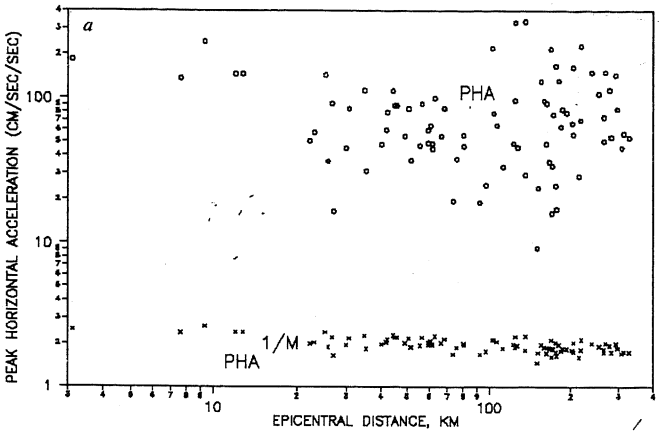


Figure 1 *a*. Observed peak horizontal acceleration (PHA) and its raised power ($PHA^{1/M}$).

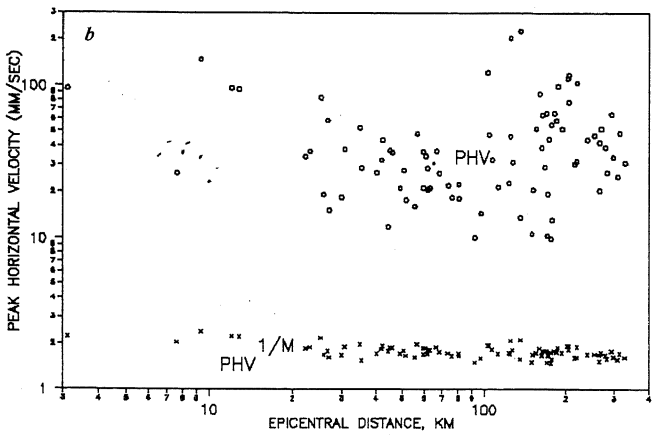


Figure 1 *b*. Observed peak horizontal velocity (PHV) and its raised power ($PHV^{1/M}$).

tion of earthquake design parameters. In order to predict peak ground motion data, the knowledge of attenuation relations of strong seismic ground motion of different regions is important. Such relations have been developed for USA[2–9], Canada[10], Japan[11] and Europe[12,13]. Such data are useful to seismologists and earthquake engineers to characterize strong ground motion and to prepare new regional seismic zoning map of different regions. In the past, most of the workers have widely used the following empirical relations given by Kanai[14]

$$\log_{10}(\mathrm{AGM}) = b_1 + b_2\ M - b_3\ \log_{10} R,$$

where AGM refers to the amplitude of ground motion parameter (acceleration or velocity) under consideration, M is earthquake magnitude, R is epicentral distance, and b_1, b_2 and b_3 are coefficients which are dependent on numerous parameters, including geology of the region.

The strong ground motion data for Indian region were almost nonexistent before 1986 and therefore no estimate for either ground motion parameters or these three coefficients was made for Indian region. After 1986, however, a number of strong motion records are available from earthquakes in the Himalayan region. In the present work, strong ground motion data[15] from five earthquakes recorded along the Himalayan region during the period 1986–91 have been examined. The details of these earthquakes are given in Table 1. The strong ground motion data have been used to determine empirical relation to predict peak horizontal acceleration and velocity for future earthquakes along the Himalayan region. Figure 1 *a* and *b* shows peak horizontal acceleration and velocity respectively, from five earthquakes that occurred along the Himalayan region (Table 1). The observed peak horizontal acceleration and velocity have been raised to power $1/M$, where M is the magnitude of the earthquake (Figure 1). Figure 1 *a* and *b* shows variations of $a^{1/M}$ and $v^{1/M}$ with epicentral distance, R, which shows almost a linear trend. From these data, we have found the following empirical relations which fit well with the observed peak horizontal acceleration (a) and velocity (v):

$$\log(a^{1/M}) = 0.433 - 0.073 \log(\mathrm{R}) \pm 0.037P, \quad (1)$$

$$\log(v^{1/M}) = 0.351 - 0.053 \log(\mathrm{R}) \pm 0.036P, \quad (2)$$

where $P = 0$ for 50th percentile and $P = 1$ for 84th percentile, and $\pm$ sign shows the upper and lower limits of the peak horizontal acceleration and velocity.

Variation of peak horizontal acceleration and velocity in Figures 2 *a–e* and 3 *a–e* is shown respectively, with epicentral distance R for all the five earthquakes that occurred along the Himalayan region. It is clear from these figures that the peak horizontal acceleration and velocity decrease with epicentral distances at different rates for different earthquakes. The amplitude of the observed peak horizontal acceleration and velocity shows variation with the epicentral distance. These variations are due to the difference in the magnitude of the individual earthquakes. The qualitative nature of the variation with the epicentral distance in all these earthquakes is almost similar. The data are very much scattered. The error bars show the 84th percentile values of the peak horizontal acceleration and velocity given for individual earthquakes (Figures 2 *a–e* and 3 *a–e*).

Using equations (1) and (2), we have computed peak horizontal acceleration and velocity for each individual

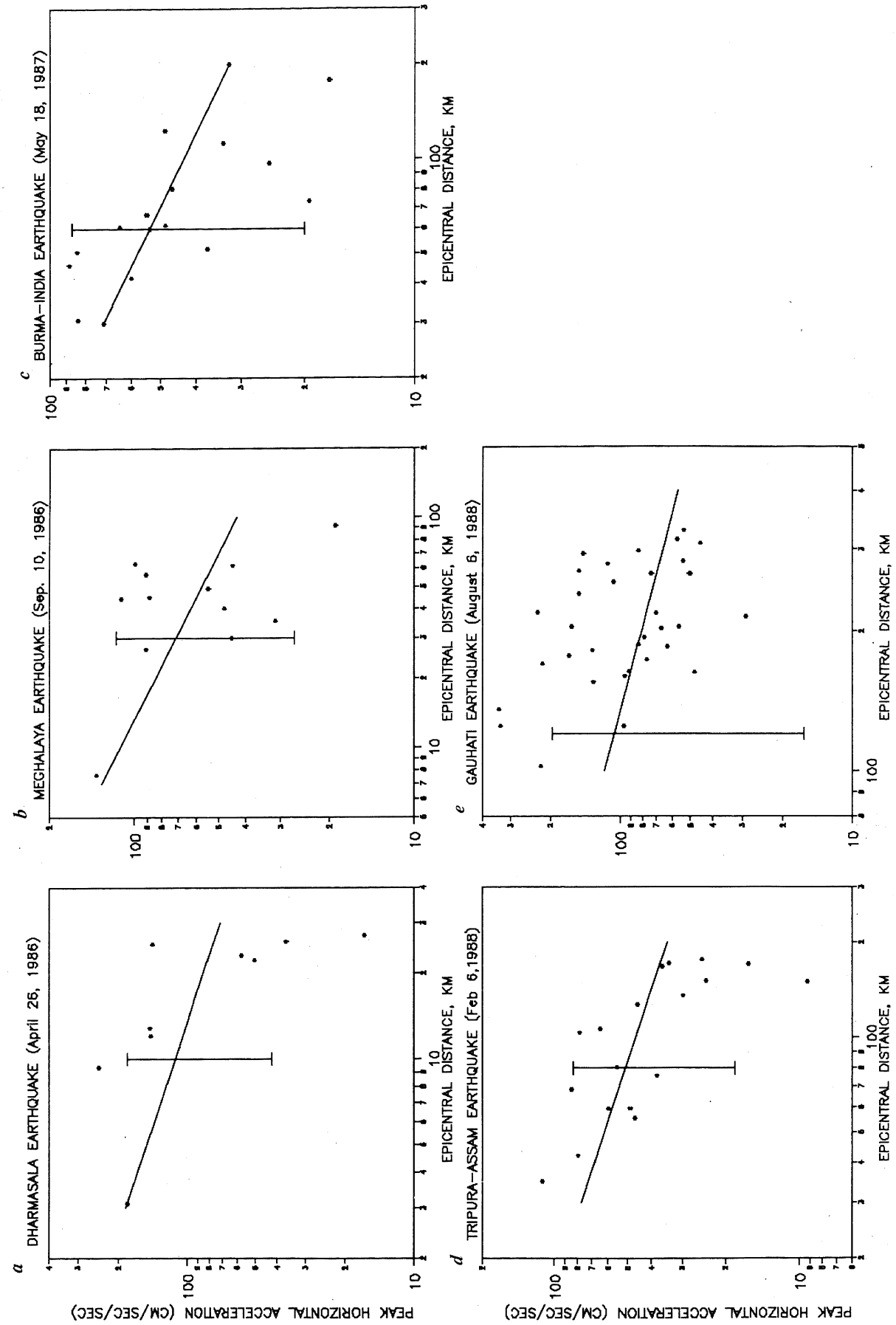


Figure 2 *a–e.* Observed peak horizontal acceleration, the solid line shows calculated acceleration from the empirical relation developed in the present work.

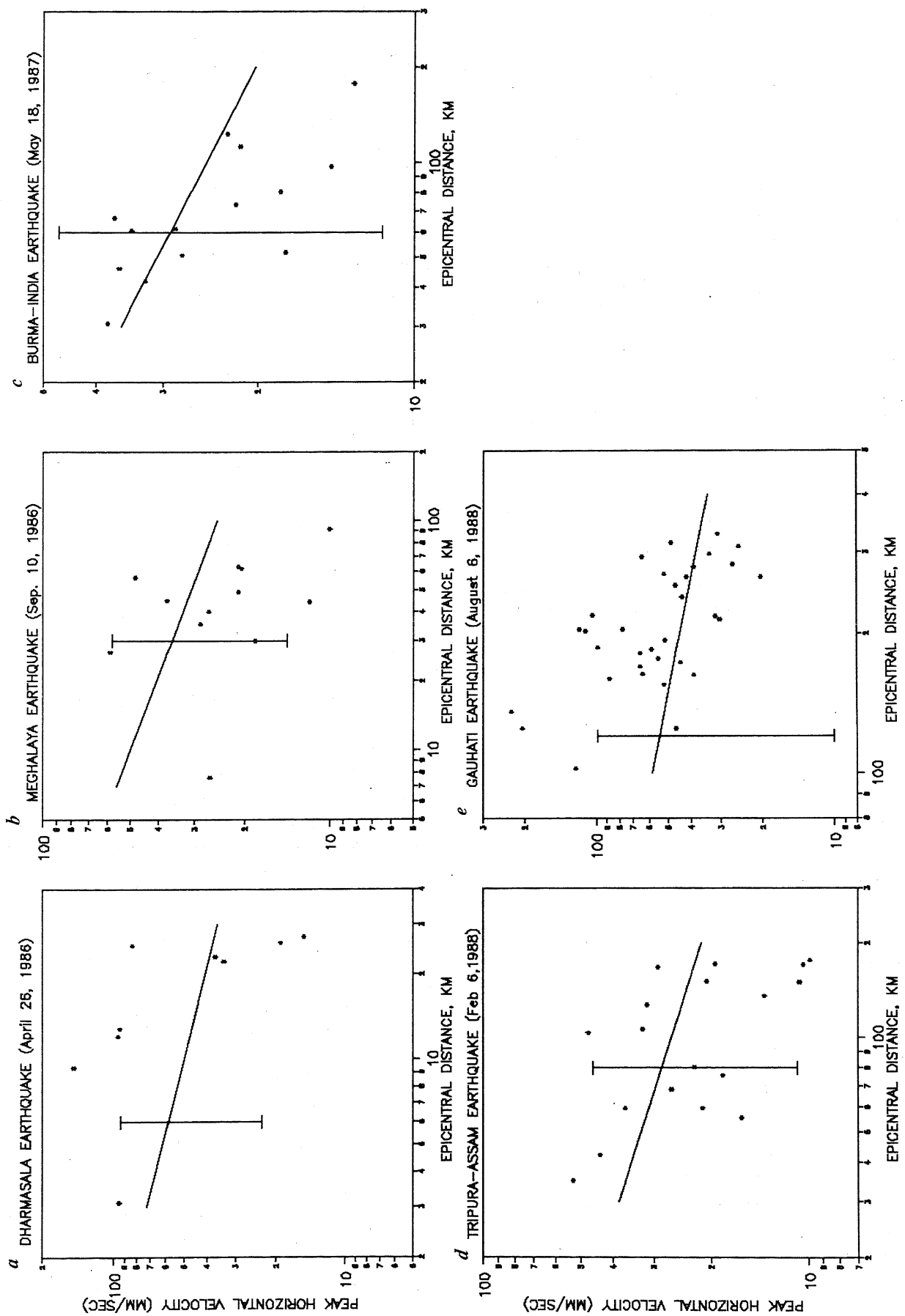


Figure 3 *a–e*. Observed peak horizontal velocity, the solid line shows calculated acceleration from the empirical relation developed in the present work.

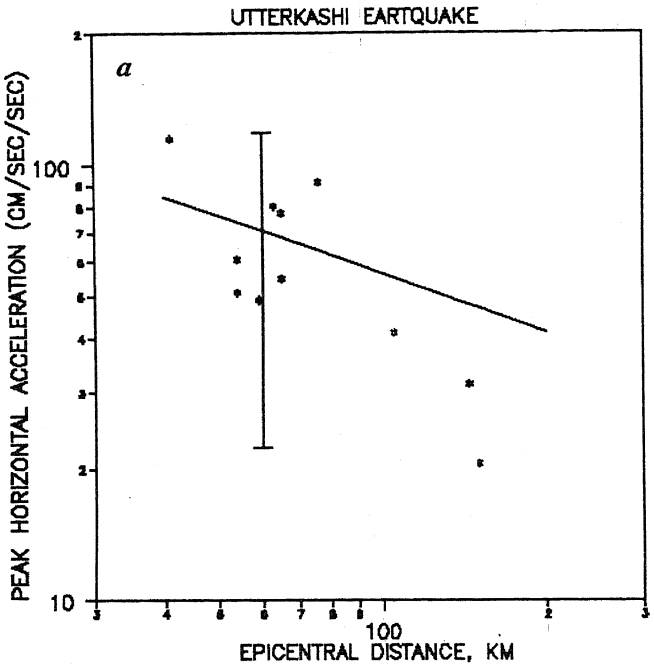


Figure 4 *a*. Observed peak acceleration in the Uttarkashi earthquake of 20 October 1991, the solid line shows theoretical predicted peak horizontal acceleration.

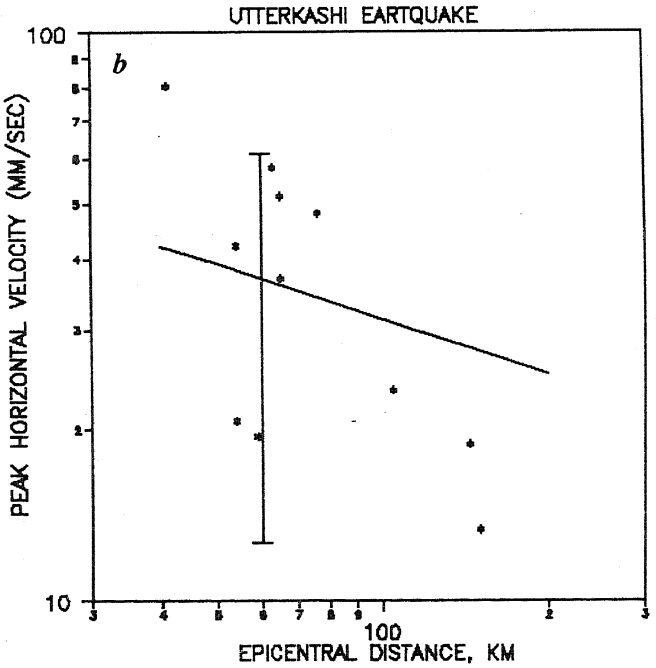


Figure 4 *b*. Observed peak velocity in the Uttarkashi earthquake of 20 October 1991, the solid line shows theoretical predicted peak horizontal velocity.

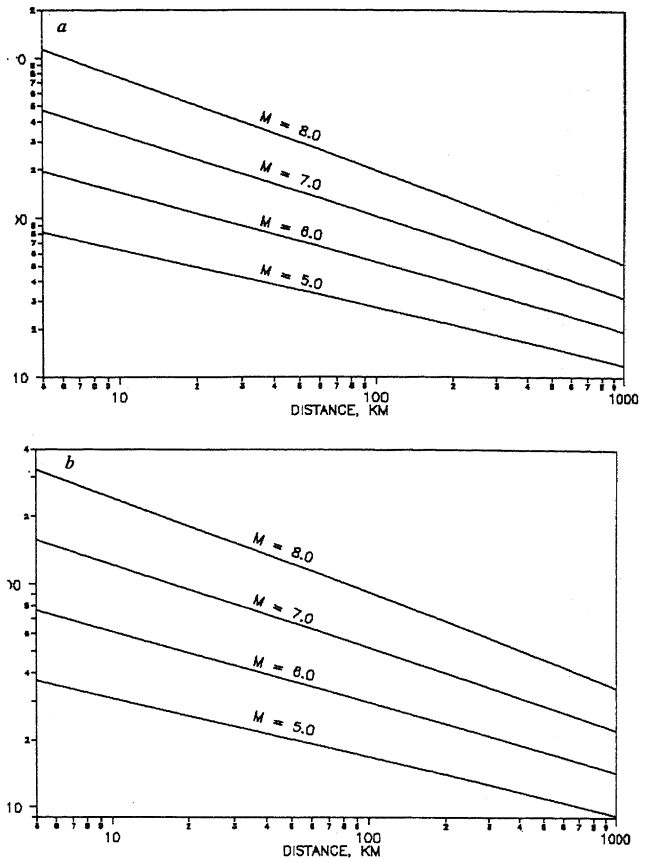


Figure 5 *a* **and** *b*. Predicted peak horizontal acceleration (*a*) and velocity (*b*) for earthquakes having various magnitudes.

earthquake which is depicted by the solid line in Figures 2 *a–e* and 3 *a–e*. It is evident that our empirical relations can explain the observed data quite nicely.

The validation of the above empirical relations has been carried out with the observed strong ground motion data[16] in the Uttarkashi earthquake of 20 October 1991 (magnitude 6.1). The observed peak horizontal acceleration and velocity of the Uttarkashi earthquake with epicentral distances are shown respectively in Figure 4 *a* and *b*. The solid curve is the predicted acceleration and velocity which is found to be in agreement with the observed strong ground motion.

Following the attenuation relations developed in the present work, we have evaluated the expected peak horizontal acceleration and velocity for earthquakes of varying magnitudes ($M = 5$, 6, 7 and 8) in the Himalayan region (Figure 5 *a* and *b*). It has been found that the maximum peak horizontal acceleration of about 1.12 *g* (Figure 5 *a*) may be observed in the vicinity of 5 km from the epicentre, which can produce peak horizontal velocity of about 320 mm/s (Figure 5 *b*) in any future earthquake in the Himalayan region.

In conclusion, we have found new empirical relation for peak horizontal acceleration and velocity. These

attenuation relations will be extremely useful in estimating strong ground motion parameters for future earthquakes in the Himalayan region. The present relations show the estimate of peak horizontal acceleration and velocity up to 1000 km from the epicentre of the earthquakes. Such relations are also needed for other major geological regions of India, e.g. Indo-gangetic basin and Peninsular regions, which can only be developed after having huge data base of the strong ground motion data from future earthquakes in these regions. Such relations will provide a useful guideline to the professional engineers while planning for major structures in the Himalayan region.

1. Newmark, N. M. and Hall, W. J., Proceedings of the Fourth World Conference on Earthquake Engineering, 1969, vol. 2, pp. 37–50.
2. Orphal, D. L. and Lahoud, J. A., *Bull. Seismol. Soc. Am.*, 1974, **64**, 1563–1574.
3. Trifunac, M. D. and Lee, V. W., *Soil Dyn. Earthq. Engg.*, 1992, **11**, 101–110.
4. Anderson, J. G., *Bull. Seismol. Soc. Am.*, 1978, **68**, 1147–1179.
5. Boore, D. M., *Bull. Seismol. Soc. Am.*, 1983, **73**, 1865–1894.
6. Campbell, K: W., *Bull. Seismol. Soc. Am.*, 1981, **71**, 2039–2070.
7. Campbell, K. W., *Bull. Seismol. Soc. Am.*, 1989, **79**, 1311–1346.
8. Joyner, W. B. and Boore, D. M., *Bull. Seismol. Soc. Am.*, 1981, **71**, 2011–2038.
9. Joyner, W. B. and Boore, D. M., *Geotech. News*, 1991, **9**, 21–26.
10. Hasegawa, H. S., Basham, P. W. and Berry, M. J., *Bull. Seismol. Soc. Am.*, 1981, **71**, 1943–1962.
11. Fukushima, Y. and Tanaka, T., *Bull. Seismol. Soc. Am.*, 1990, **80**, 757–783.
12. Murphy, J. R. and O'Brien, L. J., *Bull. Seismol. Soc. Am.*, 1977, **67**, 877–915.
13. Ambraseys, N. N. and Bommer, J. J., *Earthq. Eng. Struct. Dyn.*, 1991, **20**, 1179–1202.
14. Kanai, K., *Bull. Earthquake Res. Inst.*, 1961, **39**, 85–95.
15. Chandrasekharan, A. R. and Das, J. D., *Indian Soc. Earthq. Technol.*, 1990, **27**, 1–66.
16. Chandrasekharan, A. R. and Das, J. D., *Indian Soc. Earthq. Technol.*, 1992, **29**, 35–55.

Received 10 May 1995; revised accepted 7 October 1995

Involvement of calcium in brassinolide and auxin-induced cell elongation

Anitha Naren, T. G. Prasad, V. R. Sashidhar and M. Udaya Kumar

Department of Crop Physiology, University of Agricultural Sciences, GKVK, Bangalore 560 065, India

Coleoptile segments incubated in brassinolide (BR) or auxin (IAA) along with calmodulin inhibitor, chlorpromazine (CPZ) or calcium chelating agent (EGTA) or calcium surface antagonist, lanthanum chloride ($La^{+3}Cl_3$) resulted in inhibition of BR and IAA-induced elongation growth of coleoptiles. Pretreatment of coleoptile segments for 4 h with EGTA or $La^{+3}Cl_3$ also inhibited both BR and IAA-induced growth. Coleoptile segments with enriched endogenous levels of free calcium showed more elongation in the presence of BR or IAA. However, BR and IAA were ineffective in inducing coleoptile growth in the presence of calcium inhibitors. Results indicated that BR and IAA act through calcium and or calmodulin protein in inducing elongation growth in coleoptile.

BRASSINOSTEROID, a novel plant growth-promoting steroidal lactone, was first isolated and identified from pollen grains of brassica (*Brassica napus*)[1]. Recently, the presence of this new class of plant hormones has been shown in a number of plant species. It has also been shown to be involved in growth responses in many test systems[2–5], especially inducing cell elongation and cell division[1,5–7]. In addition to acting independently, it is found to act synergistically with other growth hormones found in plants[3,8,9]. One such interaction widely reported is that between BR and auxin in inducing cell elongation[3,8,10]. Auxin-induced cell elongation has been shown earlier to be mediated through the release of Ca^{2+} from membrane vesicles, which on complexing with the calcium-modulated protein, calmodulin, results in the response. It is known that calmodulin modulates auxin-induced growth[11]. Though the involvement of calcium in auxin action has been recognized for many years[12], the nature of interaction between auxin and calcium is poorly understood. Auxin-induced elongation in pea epicotyl is blocked by treating the segments with calcium chelators (EGTA and CTC), calcium surface antagonist (Lanthanum) and calcium channel blocker, Verapamil. The existing evidence suggests that auxin causes calcium efflux. In recent years, Ca^{2+} has been regarded as a second messenger and in conjunction with the calcium-binding protein, calmodulin, has been reported to be involved in the regulation of a number of cell processes, both in animals and plants[13–17]. Because of the close mimicking of the action of auxin by BR, we hypothesize that the mechanism underlying BR-induced cell elongation could also be mediated through calcium. We test this hypothesis here by using surface calcium antagonist, calcium chelating agent and finally potent antagonists of calcium calmodulin protein.

Brasinolide (BR) chemical was obtained from Godrej Soaps Ltd, Bombay. Calmodulin antagonist – chlorpromazine (CPZ), calcium chelating agent – ethylene glycol-*bis*(aminoethyl)tetraacetic acid (EGTA) and calcium surface antagonist – lanthanum chloride ($La^{+3}Cl_3$) and indole-3-acetic acid (IAA) were from Sigma Co, USA.

The assay of cell elongation induced by BR and IAA was carried out using wheat coleoptiles. Wheat seeds

were sown in plastic trays filled with vermiculite and kept in the dark for 5 days at 24°C. The dark grown coleoptiles were then harvested under dim red light and 10 mm segments were excised by excluding 3 mm tip using a coleoptile cutter. Excised coleoptile segments were transferred to a beaker containing distilled water and kept floating for an hour prior to the experimental treatments.

A set of 10 presoaked coleoptiles was transferred to petri dishes (5 cm diameter) containing 10 ml of incubating medium consisting of 1 mM phosphate buffer (pH 6), various calcium chelator/calcium surface antagonist solutions (CPZ, EGTA and $La^{+3}Cl_3$) along with 10 µM IAA or 0.2 µM BR; suitable controls without IAA or BR were maintained in all experiments. The coleoptiles were incubated in dark for 20 h at 24°C and the increase in the coleoptile length was measured using a millimeter scale and magnifying lens.

In another set of experiments, coleoptile segments were preincubated in the above-mentioned calcium chelators/calcium surface antagonists solution for 4 h, the coleoptile sections were rinsed in distilled water, surface-dried and then transferred to IAA or BR solutions. The response in terms of coleoptile segment elongation was monitored.

An attempt was also made to examine the effect of IAA or BR on cell elongation of coleoptiles enriched with calcium. Calcium-enriched coleoptiles were obtained by planting wheat seeds previously soaked in 1% $CaCl_2$ solution and sown in vermiculite irrigated with 1% $CaCl_2$ solution. Following germination for 5 days, the coleoptiles were processed as mentioned above and incubated in either water, IAA or BR and the increase in length was monitored.

The presence of calcium calmodulin antagonist-CPZ along with IAA or BR inhibited elongation growth of the coleoptile segments. A concentration-dependent reduction in IAA or BR-induced growth was observed with increasing concentration of CPZ in the medium. At a concentration of 75 µM of CPZ about 50% inhibition of IAA and BR-induced elongation was observed and at a concentration of 140 µM, both IAA as well as BR-induced growth were completely inhibited (Figure 1).

The addition of increasing concentrations of calcium chelating agent (EGTA) or calcium surface antagonist, $La^{+3}Cl_3$ to the incubation media progressively decreased IAA as well as BR-induced elongation. IAA and BR-induced elongation was inhibited to an extent of 50% in the presence of 2–3 mM EGTA (Figure 2) and 1–5 nM of $La^{+3}Cl_3$ (Figure 3).

Coleoptile segments pretreated with CPZ showed reduced growth when transferred to IAA and BR. There was a concentration-dependent reduction in coleoptile growth with increasing concentration of CPZ used for pretreatment. Segments incubated in IAA (10 µM)

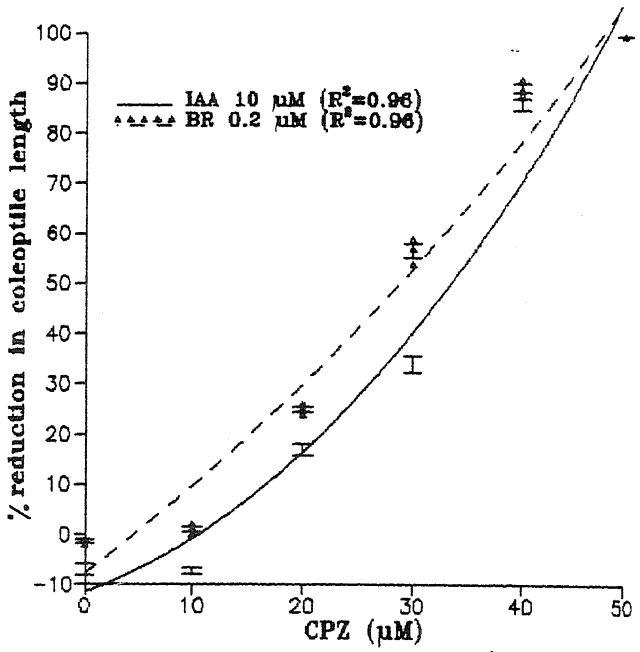


Figure 1. Effect of CPZ on IAA and BR-induced growth of wheat coleoptile segments. (Coleoptile segments were incubated in medium containing IAA (10 µM) or BR (0.2 µM) with different concentrations of CPZ. The final lengths of coleoptile segments were measured at the end of 24 h incubation in dark).

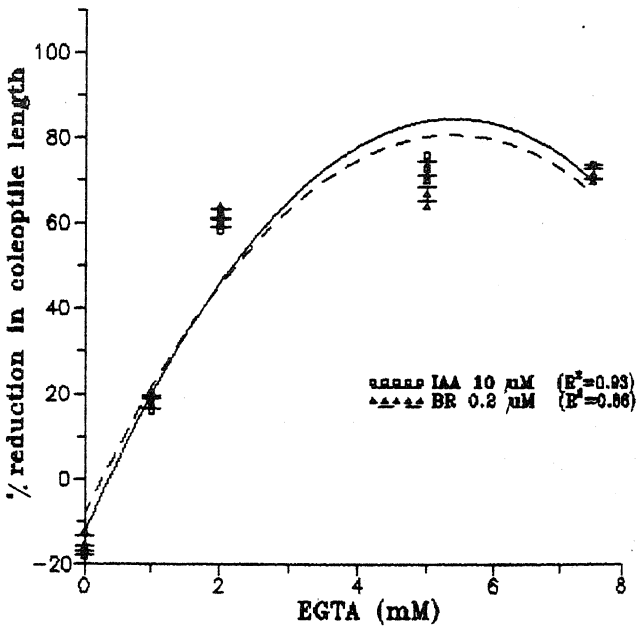


Figure 2. Effect of EGTA on IAA and BR-induced growth of wheat coleoptile segments. (Coleoptile segments were incubated in medium containing IAA (10 µM) or BR (0.2 µM) with different concentrations of EGTA. The final lengths of coleoptile segments were measured at the end of 24 h of incubation in dark).

Table 1. Effect of CPZ and EGTA on the extent of inhibition of growth of normal and calcium-enriched wheat coleoptile segments

Inhibitors	Calcium level in the coleoptile	Basal medium	IAA (10 μM)	BR (0.2 μM)
CPZ (50 μM)	Normal	30.60	81.30	63.93
	Enriched	84.33	92.16	81.00
EGTA (5 mM)	Normal	20.25	82.07	54.42
	Enriched	36.84	92.43	79.59

(% inhibition of growth over the respective control).

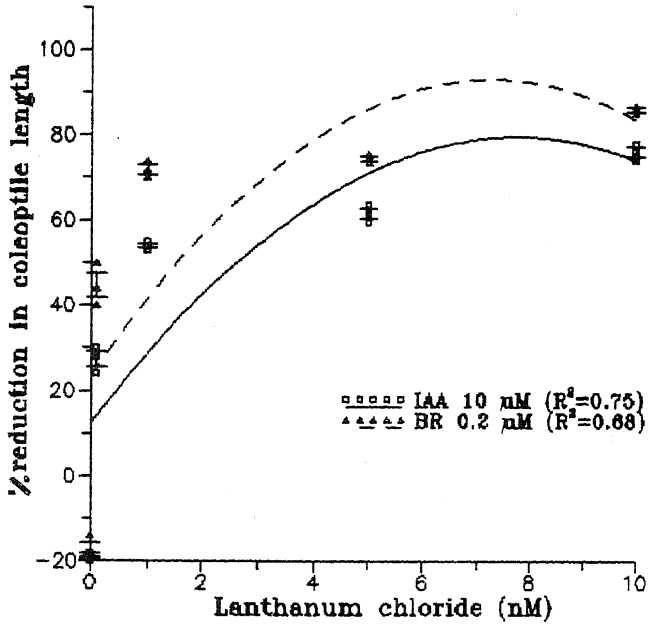


Figure 3. Effect of lanthanum chloride on IAA and BR-induced growth of wheat coleoptile segments. (Coleoptile segments were incubated in medium containing IAA (10 μM) or BR (0.2 μM) with different concentrations of $LaCl_3$. The final lengths of coleoptile segments were measured at the end of 24 h incubation in dark).

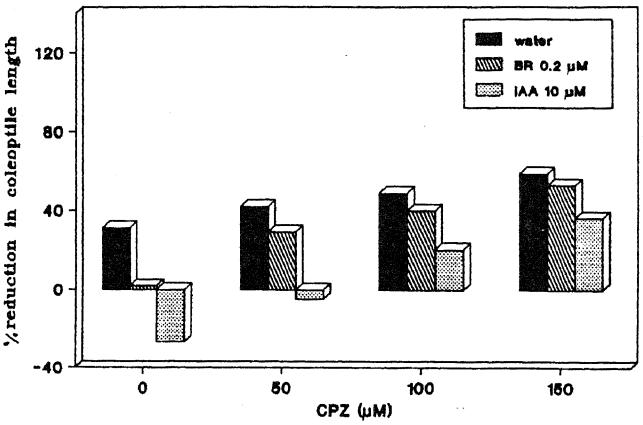


Figure 4. Effect of pretreatment of coleoptile segments with CPZ on IAA and BR-induced growth.

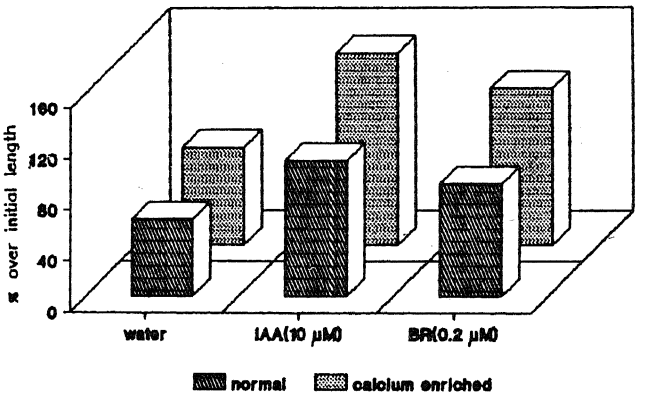


Figure 5. Effect of IAA and BR on the growth of coleoptile enriched with calcium.

medium showed 133% increase in growth. Segments pretreated with 50, 100, 150 μM of CPZ, however, showed increases of 105, 80, 64% respectively. A similar decline in response was observed when the CPZ-pretreated coleoptile segments were transferred to BR (0.2 μM) medium (Figure 4). Similar results were also obtained by pretreating coleoptiles with calcium chelating agent EGTA or calcium surface antagonist $La^{+3}Cl_3$ (data not shown).

Coleoptiles from seedlings raised in calcium-enriched medium showed significantly more elongation than coleoptile segments not enriched with calcium (Figure 5). The endogenous-free calcium in the former was nearly 27% more than the latter. Coleoptile segments enriched with calcium showed 16, 45 and 45% increase in length in control, IAA (10 μM) and BR (0.2 μM) media respectively over coleoptile segments not enriched with calcium. Under normal and enriched levels of calcium, both CPZ and EGTA were effective in reducing the elongation growth of coleoptiles induced by IAA and BR. At 50 μM CPZ and 5 mM EGTA, complete inhibition of auxin and BR-induced elongation was noticed.

Calmodulin inhibitor (CPZ) or calcium chelator (EGTA) or surface calcium antagonist ($La^{+3}Cl_3$) reduced both auxin as well as BR-induced elongation growth of coleoptiles. Coleoptiles pretreated with calcium calmodulin inhibitor – CPZ, calcium sequestering agent and calcium channel blocker significantly reduced growth in IAA or BR medium. This indicated that reducing cytoplasm-free calcium concentration and inhibiting calmodulin activity reduces IAA as well as BR-induced growth.

Coleoptile segments with high endogenous calcium concentration showed more elongation growth in the presence of BR or IAA in the medium. All the calcium and calmodulin inhibitors used were also effective in inhibiting IAA or BR-stimulated growth of coleoptile

tissue with elevated calcium levels (Table 1). The results emphasize two points—the IAA- and BR-induced growth was more in tissues having high cytosolic-free calcium concentration and calcium inhibitors were effective in reducing the hormone-induced growth in tissues with elevated calcium concentration. This may be due to altered calmodulin content in the cytoplasm. It is opined that altered calcium levels may influence calmodulin activity and thus calmodulin-mediated cellular processes[18,19]. The range of concentration of calcium inhibitors which resulted in significant reduction in growth was comparable to the inhibition of other calcium or calmodulin-mediated responses[20].

Our results therefore indicate that the action of BR which closely mimics that of auxin in causing cell elongation could also be mediated by the release of calcium with the involvement of the calmodulin-binding protein.

1. Mitchell, J. W., Mandava, N., Worley, J. F., Plimmer, J. R. and Smith, M. V., *Nature*, 1970, **225**, 1065.
2. Wada, K., Maruma, S., Ikekawa, N., Morisaki, M. and Mori, K., *Plant Cell Physiol.*, 1982, **22**, 323.
3. Sasse, J. M., *Physiol. Plant.*, 1990, **80**, 401–408.
4. Gregory, J. F. and Mandava, N. B., *Physiol. Plant.*, 1982, **54**, 239.
5. Worley, J. F. and Mitchell, J. W., *J. Am. Soc. Hortic. Sci.*, 1971, **96**, 270.
6. Mitchell, J. W. and Gregory, L. E., *Nature*, 1972, **239**, 253–254.
7. Gregory, L. E., *Am. J. Bot.*, 1981, **68**, 586–588.
8. Yopp, J. H., Mandava, N. B. and Sasse, J. M., *Physiol. Plant.*, 1981, **53**, 445.
9. Katsumi. M., *Plant Cell Physiol.*, 1985, **26**, 615–625.
10. Wada, K., Maruma, S., Abe, H., Morishita, T., Nakamura, K., Uchiyama, M. and Mori, K., *Agric. Biol. Chem.*, 1984, **48**, 719.
11. Ragothama, K. G., Mizrahi, Y. and Poovaiah, B. W., *Plant Physiol.*, 1985, **79**, 28–33.
12. Cleland, R. E. and Rayle, D. L., *Plant Physiol.*, 1977, **60**, 709–712.
13. Anderson, J. N. and Cormier, M. J., *Biochem. Biophys. Res. Commun.*, 1978, **84**, 595–602.
14. Cormier, M. J., Jarett, H. W. and Charbonneau, H., in *Calmodulin and Intracellular Receptors* (eds Kakiuchi, S., Hidaka, H. and Means, A. R.), Plenum, New York, 1982, pp. 125–139.
15. Cheung, W. Y., *Science*, 1980, **207**, 19–27.
16. Poovaiah, B. W. and Leopald, A. C., *Plant Physiol.*, 1974, **54**, 289–293.
17. Roux, S. J. and Stocum, R. D., in *Calcium and Cell Function* (ed Cheung, W. Y.), Academic Press, New York, 1983, vol. 3, pp. 409–453.
18. Paliyath, G. and Poovaiah, B. W., *Proc. Natl. Acad. Sci. USA*, 1984, **81**, 2065–2069.
19. Sasaki, Y. and Hidaka, H., *Biochem. Biophys. Res. Commun.*, 1982, **104**, 451–456.
20. Dieter, P. and Marme, D., *FEBS Lett.*, 1981, **125**, 245–248.

ACKNOWLEDGEMENTS. We thank Dr Vyas and Dr P. Rama Raj from Godrej Soaps Ltd for their encouragement. This work was supported by a collaborative research grant provided by Godrej Soaps Ltd, Bombay.

Received 22 June 1995; revised accepted 26 September 1995

Uttarkashi Earthquake. H. K. Gupta and G. D. Gupta, eds. Geological Society of India, Bangalore. 1995. xxii + 233 pp.

The Uttarkashi earthquake of 19 October 1991 was important in many ways. Firstly, this was the first earthquake of magnitude (M_s) 7.0 in the Indian domain for quite a number of years. Secondly, the casualty figure was high and the damage to dwellings excessive. The loss occurred in an already backward region of India. Thirdly, this was the first substantial earthquake in the vicinity of the Tehri dam site since the designs for the Tehri hydroelectric and irrigation project were completed in the late nineteen sixties. Fourthly, the earthquake triggered a number of accelerographs installed in the region. These records are valuable evidence in the debate about seismic hazards to the Tehri project. Fifthly, the first Indian estimate of the epicentral location for the earthquake was in error grossly. The earthquake also exposed the lack of preparedness of the civil administration for relief and succour operations.

A conference on the Uttarkashi earthquake was organized at New Delhi in November 1992. The volume under review is a compilation of 19 articles presented at the conference.

Chandrasekaran and J. Das report analyses of records from 13 US-made accelerographs that were triggered by the Uttarkashi earthquake in the Garhwal region. The maximum peak ground acceleration was 0.30 *g* and 0.29 *g* at Bhatwari and Uttarkashi respectively, the two stations closest to the epicentre. Brijesh Chandra and others in a companion article report recorded peak ground acceleration of 0.52 *g* at Uttarkashi where an Indian-made accelerograph was operating also. Both the articles present information of interest to earthquake engineers, civil engineers and architects. But there should have been a more critical examination of the data in these articles. Since they are from the same university department, cross-referencing would have been possible and it would have been useful to the readers. A question of great importance left unanswered nonchalantly is, Why did the Indian-made instruments show higher accelerations than the imported instruments at the same station?

S. Jain and S. Das analysed the accelerograms reported in the above-mentioned article by Chandrasekaran and Das to assess the provisions of the Code IS 1893: 1984 for design of earthquake-resistant structures in different parts of India. It may be recalled that India has been divided into zones I to V in the code. Maximum attention has to be paid in design of structures for parts of India assigned to zone V. The authors conclude that provisions of the code are conservative for zones I to III. Thus, the cost of important structures with provisions for resistance to earthquake forces could be lowered over most parts of India if the codal provisions are relaxed suitably. On the other hand, the existing provisions are not sufficiently conservative for zone IV. This is crucial because Tehri dam site has been assigned to zone IV. In simple terms, if the observation of Jain and Das is applied to Tehri dam, then it follows that the dam built for earthquake resistance according to guidelines for zone IV may be exposed to greater earthquake hazards than those for which it would have been designed. The Uttarkashi earthquake did not provide scope for assessing the codal provisions for zone V.

Two articles by Roy and Hasija deal with geodetic observations. Elevations of bench-marks between Saharanpur and Mussoorie via Dehra Dun measured in 1991 and 1992, before and after the Uttarkashi earthquake, are compared with those measured in 1974–77 in the first article. It appears that bench-marks between Dehra Dun and Mussoorie experienced subsidence of up to 7 cm between 1991 and 1992. The authors attribute the subsidence to the Uttarkashi earthquake. But this conclusion is difficult to accept in the absence of supporting simulations based on the elastic dislocation theory. The bench-marks in question are located 70 to 80 km from the epicentre and the distance may be too large for the observed magnitude of subsidence. The suggestion of the authors that the bench-marks should be occupied regularly should be heeded by the authorities.

The second article deals with geodetic observations elsewhere in the Garhwal Himalaya prior to the Uttarkashi earthquake. This article is a wasted effort scientifically. The dimensions of the survey figures measured were too small. The geographic locations of the figures have been masked. Thus we are not sure about the region of investigation. The authors belong to Survey of India whose *raison d'etre* is to determine and disseminate accurate locations of points in the Indian territory. But the maps appearing in the two articles do not have latitude and longitude information.

Dattatrayam and others report analysis of digital data recorded through a seismograph at New Delhi for three aftershocks of the Uttarkashi earthquake. The estimated seismic moments are of the order of 10^{13} Nm and stress drops are in the range of 0.2 to 0.44 MPa. The authors present similar results for two earthquakes occurring 5 and 3 days prior to the Uttarkashi earthquake. They call them foreshocks. But their reported epicentres are so far from the epicentre of the main earthquake that this appellation is unacceptable. The low stress drops deserve note and further scrutiny.

There is an article by Arya entitled 'Long term measures for earthquake protection in Uttranchal Area of Uttar Pradesh, India'. The theme has been exposed by the author earlier for the Dharamsala area of Himachal Pradesh. Arya presents estimates of numbers of houses that may collapse or be damaged if an earthquake of magnitude (*sic*) of 6.5 occurs with its epicentre in one of the eight administrative districts in the hills of west Uttar Pradesh. Arya estimates that Almora and Dehra Dun districts will have maximum numbers of collapsed and damaged houses respectively. It is a pity that he did not check his predictions with the estimates of houses actually damaged or collapsed during the Uttarkashi earthquake.

The later part of the article provides sane advice for the future. Thus, all future constructions in these hill regions should have earthquake-resistant features. The additional cost in the case of the most cheaply constructed houses is estimated to be about 10% of the total cost. Arya notes that strengthing of existing houses is possible but will incur higher percentage of original cost.

Conditions in the affected region immediately after the Uttarkashi earthquake were tough. Efforts of scientists who visited the area for observations in this period deserve appreciation. Still the fact remains that the articles which describe the results of these qualitative and mildly quantitative observations are of a routine nature. Many articles contain conjectures about the origin of the earthquake. Extrapolations from surface geological obser-

vations to hypocentral depths of 10 to 12 km appear unwarranted.

Perusal of the articles of this volume collectively reveals in graphic detail the relatively low status of seismological studies in India *circa* 1991. Some mitigating factors may be cited. Lack of funds for seismological observations and research on a scale commensurate with the vast size of our country would be at the top of the list. The situation has not improved since 1991 at least in so far as seismological observations in our northern hill regions are concerned. Similarly, progress towards incorporation of earthquake-resistant features in individual houses and dwellings is abysmal. Thus the portents are clear. The death and destruction during the Uttarkashi earthquake will be repeated in all its tragic details during the next major earthquake of the Himalaya.

RAMESH CHANDER

Department of Earth Sciences,
University of Roorkee,
Roorkee 247 667, India

MEETINGS/SYMPOSIA/SEMINARS

IAMAS/IAPSO Joint Scientific Assemblies

Date: 1–9 July 1997
Place: World Congress Centre, Melbourne, Australia

Joint assemblies of the International Association of Meteorology and Atmospheric Sciences (IAMAS) and the International Association for the Physical Sciences of the Oceans (IAPSO), two of the associations constituting the International Union of the Geodesy and Geophysics (IUGG) will be held at Australia. In addition, there will be a symposium sponsored by the Tsunami Commission of the IUGG and IUGG associations.

Contact: IAMAS/IAPSO Secretariat
Convention Network
224 Rouse Street, Port Melbourne Victoria 3207
Australia
Phone: +613 9646 4122
Fax: +61 3 9646 7737
E-mail: mscarlett @peg.apc.org

11th Himalayan – Karakoram – Tibet Workshop

Date: 28 April–1 May 1996
Place: Flagstaff, Arizona, USA

There will be three special sessions, Geodynamic Models of the Himalayas and Tibet, Neotectonics and Quaternary Geology of the Himalayas and Tibet and Extensional Tectonics in a Compressional Orogenic System.

Contact: Prof. Allison Macfarlane
Dept. of Geography & Earth System Science,
George Mason University, 1E2, Fairfax,
VA 22030-4444
USA
Phone: 703/993-1207
Fax: 703/993-1216
E-mail: amacfarl @osf1.gmu.edu

Sixth All India Conference on Cytology and Genetics

Date: 23–25 February 1996
Place: Rohtak

Topics include: Animal cytology and genetics, plant cytology and genetics, cell biology and molecular genetics, genetics and plant breeding, mutagenesis, population genetics, evolutionary genetics and biosystematics.

Contact: Prof. Ravi Prakash
Local Secretary
Department of Biosciences
Maharshi Dayanand University
Rohtak 124 001

DAE Solid State Physics Symposium

Date: 27–31 December 1995
Place: Calcutta

Topics include: Phonon physics, Electron states and electronic properties, Magnetism and magnetic properties, Semiconductor physics, Physics of defects and disordered materials, Transport properties, Superconductivity and superfluidity, Liquids, liquid crystals and plastic crystals, Phase transitions and critical phenomena, Surface and interface physics, Non-equilibrium phenomena in solids, Physics of complex systems, Resonance studies and relaxation phenomena, Solid state devices, techniques and instrumentation.

Contact: Shri Saibal Basu
Secretary, Solid State Physics Symposium
Solid State Physics Division
Bhabha Atomic Research Centre
Bombay 400 085
Phone: (022) 5563060
Fax: 91-22-5560750, 91-22-5560534
E-mail: sps @ magnum.barct1.ernet.in

In the pursuit of science: The 'subjective' element*

A. N. Mitra

Several facets of 'subjectivity' in science, from abstract value judgement to mundane human level interactions, are discussed vis-à-vis its traditional 'objective' image.

There is an almost instinctive perception in any unsuspecting mind that the pursuit of science is necessarily an objective exercise, intensely rational in its scope, and singularly free from any 'pre-conceived' notions. Indeed, the standard scenario of an interplay of theory versus experiment for determining the outcome of any idea in science, major or minor, is a *prima facie* evidence for this basic perception. In the language of Gerald Holton[1], this 'playground' is the so-called *contingency plane*, one in which theory (the calculus of logic and mathematics) is depicted as an ordinate, and experiment (empirical matters of fact and 'data') as the abscissa. Formally, this is called the 'inductive method', one in which a five-step process is involved, according to the physicist Fredrich Dessauer[2]. These five steps are simple, and in conformity with common sense[1,3]: (i) hypothesize a provisional statement; (ii) refine and structure the hypothesis; (iii) draw logical conclusions or 'predictions' which can stand experimental check; (iv) then check the predictions against the 'data' (experiment); (v) if both tally within experimental limits, then a warrant is available for claiming a 'universal validity' for the result.

On the face of it, this looks like an airtight procedure, with hardly any scope for intrusion of any other degree of freedom. However, a little reflection would show that such a picture is a gross oversimplification. For, one may ask, what is the nature or rather the origin of the hypothesis itself on whose basis the theorizing and experimentation were planned and executed? More specifically, what is the source of the original induction[1] or idea? It is here that an 'undefined' ingredient is suspected to have (unobtrusively) entered the fray already! This ingredient is quite distinct from the two-component (theory versus experiment) characterization of the 'induction process' introduced above. It is far more nebulous than the straight-jacket theory–experiment classification, and often 'resists' a quantitative definition. And yet it is a vital component of the structure of modern science, and has all along played a vital underlying role in the shaping of scientific theories since the Kepler–Newton days. Kepler had a twofold philosophy of the Universe (by fiat?), viz. (i) a mathematical harmony on the one hand and (ii) a central theological order on the other (no 'proof' offered here!). Newton too had his five rules[1,3], of which the first four are quite 'rational' (simplicity and economy of description; uniformity of Nature; universality (no exceptions); new hypotheses disallowed against contrary evidence). However, he had a fifth rule, not explicitly stated, concerning the very definition of a hypothesis.

Holton in his book[1] calls it the thematic component, which cannot be 'tested' in the sense of theory vs experiment, yet cannot be wished away[3]. One might then wonder if such a concept goes against the standard philosophy of Karl Popper, viz. the essential characteristic of any science is the 'falsifiability' of its theories. A possible answer is that no obvious contradiction is involved, since the Popperian criterion applies only to the structured part of the hypothesis or 'themata' in the form of a 'theory' which must face confrontation against experiment, towards eventual victory or defeat. On the other hand, the 'themata' is more like the 'spirit' or soul behind its structured form (the theory), and cannot be contained within the four walls of the laboratory. It resides in the human mind: it can reassert itself, or make a comeback in another form, or disappear from the scene altogether, all depending upon the nature of the outcome of its 'present incarnation' (the structured theory) with experiment. The point is that it is basically 'free', much like the proverbial '*die Gedanken*':

'Die Gedanken sind frei; wer kann sie erraten?
Sie fliegen vorbei; es nechliche schatten.
Kein Mann kann sie wiessen; keine Yaeger erschissen;
Es bleibeth dabei; die Gedanken sind frei.'

It is this thematic element that turns out to be a subjective concept inasmuch as it resides in the human mind with all its complex dimensions, which cannot be quantified. It is mostly concerned with a sort of 'value judgement' on the 'quality of ideas' which (by experience?) are deemed relevant for the 'understanding of nature' and those which are not. For example, past experience strongly suggests that certain attributes like 'beauty' and 'naturalness' which cannot be quantified as such (but 'felt' by the mind all the same!) have all along played a vital role in shaping the course of science, but only their sufficiently structured forms, say in the form of some well-defined principles, such as the 'principle of extremum' in the area of physical sciences, have been successfully confronted with experiment. Indeed, it is even possible to classify the thematic element somewhat more systematically into different categories in accordance with their 'performance', i.e. quantitative contact with 'data', in the contingent plane[1]. In particular, three such categories have been suggested[3,4] – principles, postulates and prejudices[4] – in accordance with the nature of successful contact with data (after suitable structuring) in descending order, and the resultant 'degree of confidence' thus generated with respect to their credibility within the scientific community. In this respect principles have undergone so many acid tests through continuous feedback from experiment that their speculative or 'subjective' aspects have by and large been forgotten, and they are now regarded as an integral part of the scientific discipline concerned. Postulates, on the other hand (though highly structured), have not quite made the grade

*This is an expa[illegible] a lecture given at NIAS on 28 Apri[illegible]. I am grateful to Dr S Ramaseshan for urging me to write it. I also express my deep appreciation to Dr R. Ramanna, Director, for the warm hospitality of the Institute.

in terms of data support (e.g. supersymmetry). These topics are spelt out in more detail elsewhere[3,4] and are not the main theme of this discussion.

The subjective element: value judgement of concepts

At the other end of the spectrum lie certain themata, termed prejudices[3,4], which, despite their strong appeal to intuition, are somewhat nebulous in character: they resist an unambiguous formulation. On the other hand, their sweep seems to be even wider if anything, and confidence in their intrinsic power (despite their nebulosity) has somehow been sustained (often strengthened) due to past experience. Some famous examples in this highly subjective category are well within the theme of this discussion so as to stand some repetition[3,4]:

(i) 'Naturalness' as a selectivity criterion (impossible to define, yet easy to feel in a given context).

(ii) 'Beauty' and 'simplicity' in physical concepts (same problem).

(iii) 'Duality' in physical sciences (can be illustrated with many examples but not more generally definable)[5].

(iv) 'Bootstrap' philosophy (somewhat akin to Mach's principle; once very popular in particle physics; coming up again in the context of modern string theory[6]).

(v) 'Shell structure' *ad infinitum* (so far only three stages have been confirmed, though, according to an unpublished theorem due to Drell, any regularity checked up to three stages is valid generally; see also ref. 7).

These examples illustrate but one aspect of subjectivity which relates to a sort of 'value judgement' in the pursuit of science, and operates mostly in the abstract plane of pure thought process. This 'noble' aspect, which may be more appropriately defined as 'natural subjectivity' (God is not neutral as between good and evil; 'beauty' vs 'ugliness'; simplicity vs complexity; and Nature seems to conform to His choice!), represents a perception which is largely the result of 'collective wisdom' of the human race acquired over a vast millennium in evolving its way of life, and in the process unearthing the secrets of Nature. And in the hands of scientists too, the manner in which this 'noble' form of subjectivity (value judgement) came into being (despite its spontaneous appeal in most individual minds) is the result of a complex feedback process involving interactions at the human level, and hence subject to all factors prone to human failings and weaknesses. They often relate to various techniques of 'persuasive argumentation' (to borrow a phrase from Thomas Kuhn's famous book[8], which concerns much of the discussion to follow), with an obvious connotation of 'survival of the fittest'[8]. This (second) aspect of 'subjectivity' in science (which is not always 'noble' in its working!) is the central theme for the rest of this discussion with the help of illustrative examples.

Subjectivity: the human factor

Since this field is vast in itself, it is first necessary to delineate the scope of this discussion by ruling out certain topics at the outset. Thus, we shall not be concerned with rather esoteric philosophical issues like the nature of the physical laws themselves[9]: Do these laws have an 'objective reality' in the sense of an 'out there' existence, ready to be discovered by a Newton or an Einstein, or must one be satisfied with a Bohr-like effective description, *à la* the Copenhagen interpretation (CI)? Although such questions have always retained their time-honoured philosophical significance, their renewed interest in recent times owes its origin to fresh efforts directed towards bridging the conceptual gap between the CI and a formal understanding of some famous paradoxes through the strengthening of their conceptual foundations, with extensive reviews both technical[10] and non-technical[6] available in the literature. At the moment the upshot of all this highly intellectual exercise seems to be that while the CI, thus strengthened[10], is in a better position to take on its critics, the operative value of the new machinery in terms of a capacity to predict new observable phenomena is quite limited[11]. Nor should we be unduly exercised over a proper definition of science *vis-à-vis* pseudo-science[12], since any attempt to provide a sharp distinction is fraught with serious ambiguities[12], and instead take a pragmatic, matter-of-fact view in this regard. Instead, we shall find it more relevant to our chosen theme of 'human subjectivity' to focus on the attitudes of scientists[8] within their own community in the matter of implementing the time-honoured theory–experiment confrontation in actual practice. To that end we shall begin with recalling, again at the cost of some repetition[3], the salient points of Kuhn's thesis[8] before going on to discuss some specific examples.

Kuhn's thesis: the arbitrary element

The evolution of science is a very complex process wherein its history is not a mere collection or catalogue of facts, theories and methods. Indeed, there exist many *inter se* contradictions in them, and only a directed stream emerges at the end after 'resolving' such contradictions. It is only such 'predigested' material at the hands of the chronicler of science that goes into the textbooks. This resolution is admittedly a democratic process in the usual sense of the term, and it is here that there is a catch! How about those ideas, theories (and even experiments) that did not make the grade? Indeed, it is entirely possible that there could well have been many good points about them, but somehow they got lost out in the race, not always because they got rejected by experiment (or superior experiment), but often because they never got a chance to come up to that stage! Perhaps their theories did not have strong enough godfathers (or willing listeners) for 'marketing' them. One must hasten to add that this is by and large a fair game, but the main point of emphasis (and not infrequent concern) is that in this complex process of evolution there is a fierce element of competition (often bordering on rivalry) in which the 'fittest' of ideas is given the best chance to survive. And the definition of 'fittest' is not always reassuring, even though a minimum criterion is one of maximum internal consistency plus fits to the data. Now suppose there are more than one such candidates, all conforming to these minimum criteria: Which one should be chosen? The question is not necessarily as hypothetical as it sounds, especially if such theories are proposed at some 'branching-off' stage (or cross-roads?) in the process of evolution of the science concerned. In such cases the 'lucky one' often gets chosen in an apparently arbitrary fashion, to the detriment of the others with equal *a priori* claims. True, such a choice involves a

broad consensus within the community; yet it implies a preselection[8] at the current state of the art (not waiting for more discriminating criteria at a later date). Such a process, by its very nature, is a highly subjective exercise, and gives science a very different image from that of an 'objective pursuit' of truth.

The essential point here is that the inadequacy of methodology[8] introduces an arbitrary element into the field in the course of its evolution, through a continual competition between a number of distinct views of Nature, each partially compatible with experiment. How did this come about? Somewhere along the line a particular 'belief structure' got espoused by the community[8], to the detriment of the other claimants (the arbitrary element), due to personal or historical coincidences. A sort of spontaneous symmetry breaking occurred in the branching process wherein a particular branch has been followed[6,8]. If the methodology were more perfect, other competing ideas would perhaps have had a better chance. However, taking things as they occurred one must say that the 'chosen belief branch' at a crucial turning point in history would get strengthened with time due to its pursuit by a large and committed band of well-trained workers who have learned to work within well-defined 'conceptual boxes'[8], but somehow were not encouraged to think beyond that particular discipline. This is what constitutes normal science, whose constraints strongly inhibit any novelty that clashes with its basic tenets. This may often be broken by an 'outsider' to the field who probably 'sees' more than such constrained workers (a pointer to the benefits of interdisciplinary research?).

And yet, even within the precincts of normal science, there exists an inherent 'governor mechanism'[3] which eventually comes to the rescue, and prevents an indefinite suppression of novelty. For there appear fresh contradictions – 'anomalies' – when a normal problem fails to respond to standard rules and procedures, leading to a paradigm shift[8] – a new set of commitments or rules[3,8]. Yet the fact remains that normal science is a rather conservative process of accretion of knowledge which does not as a routine aim at novelties of fact or of theory. It likes to remain quite smug in its 'understanding' of Nature in terms of its stated rules and procedures, howsoever imperfect or inadequate they may look to its own eyes. It is only when it (normal science) is 'pushed to the corner', so to say, due to the appearance of new and unsuspected phenomena, that it feels helpless against the onslaught of fresh ideas which have a potential to lead to new theories which may even be radical.

Subjectivity: the community and the establishment

Kuhn's thesis, as summarized above, brings out the nature of subjectivity in science as seen through its actual working. There is a strong 'directional element' in the *personality* that is normal science. Within this scenario, the 'objective' aspect of science is merely confined to laboratory investigations, while the dominant role is played by its 'subjective' partner, which offers a rather 'chauvinistic' view of science as a whole, one in which the survival of the fittest is the order of the day. True, it is the laboratory investigations that provide the basic strength to its practitioners to 'fight for survival'; yet a better measure of their power lies in their debating skill, or in Kuhn's words, in their capacity for 'persuasive argumentation'[8], which is probably the dominant factor that determines their 'standing' in the scientific community, while the objective part (laboratory investigation) remains largely in the background. In this dual scenario, science appears as an organized structure governed by the establishment within the community, much like in a 'democratic socio-political entity', where the final arbiter is the 'assent of the community'[8]. Nevertheless, the establishment has its own authority which 'determines' the direction of science, despite the existence of occasional 'antiestablishment' or dissident groups within the community that are at best 'tolerated' in the sense of being 'allowed' to publish their own ideas in journals which in turn are controlled by the former. However, one cannot be too sure even on such matters, what with the information explosion on 'ideas' leading to a proliferation of 'publication aspirants', who must be kept in 'reasonable check'. In this respect the prospects of publication are better ensured with 'proximity to establishment' than through the freelance efforts of an obscure worker. As further evidence of 'centralization' around the establishment in the high-tech atmosphere of today's science, yesterday's culture of individual (small-scale) scientific enterprise is gradually yielding place to a more organized form of collective (large-scale) science, with thousand members not uncommon in an experimental team. The most recent example of such team work is provided by the spectacular manner in which the CDF and D-ZERO groups have finally 'zeroed in' on the elusive TEE quark[13].

Now while such 'big science' is by and large inevitable as well as beneficial for the planning and execution of a major project in the experimental field, it is not equally obvious that the same is also true in the *theoretical* domain, where too much 'direction' may well prove counterproductive. A topical example is provided by the amount of organized effort by theoreticians in the early eighties (under the aegis of some important centres) in predicting a finite lifetime for proton decay within the limits of experimental measurability. These limits were variously estimated by methods which paid due regard to the parameters of the (high-profile) 'high-energy' dimension of the problem (grand unification theory), without a matching concern for the (mundane) 'low-energy' (hadronic matrix element) part, which required a much greater effort to produce a physically reliable figure, even though the measured value, if any, could not possibly distinguish between the two components. An independent estimate, one in which greater attention had been paid to the hadronic part, produced a value which was more than hundred times larger than most of the above (one-sided) estimates, a limit that had not been violated by the then available data[14], but coming as it did from an obscure source far away from the (theoretical) establishment, this finding was largely ignored in the literature. This result was *not* intended to claim that the GUT theory could still be valid – perhaps not – yet it seemed to be rejected for the wrong reason. (Other examples of 'subjective' decisions arrived at prematurely by consensus are given further below.) A related trend in the same context which brings out yet another dimension of subjectivity (by way of wishing to 'conform' to the 'views' of the establishment) was also discernable around the same time, in the anxiety of certain experimental groups to appear to agree with the prediction of 'faster' rates (10^{31} years) prevalent at that time[14].

The analogy with the sociopolitical scenario could indeed be pushed further if science were to be pursued under one (world) organization, in keeping with its supposedly common (international) standard. In its actual working, however, the pursuit of science is heavily circumscribed by the local and regional environments of the scientists themselves, and this often constrains their relative opportunities *vis-à-vis* their more fortunate counterparts residing in more visible centres of scientific power. This circumstance, in turn, brings in further 'asymmetries' in the overall scientific culture, and adds yet another dimension to the 'subjectivity' factor, one governed by the *actual* (not just by analogy!) sociopolitical environments of the workers themselves.

Subjectivity: theory vs experiment

Even without pursuing the sociopolitical analogy any further, the subjective nature of the actual working of science may be brought out with some specific instances of its impact on a global scale. One example was cited just above in the context of proton decay, but it need not (indeed, should not) give the impression that subjectivity is necessarily 'bad' for the growth of science on a global scale, despite the 'local' suffering of individual workers. Yet the examples are quite instructive insofar as they help in exploding the myth about the so-called objectivity in science. In a recent book[15], Collins and Pinch have highlighted some famous and not-so-famous instances of how the subjectivity factor has prevented unwarranted fluctuations in the overall stability of basically 'correct' premises of science, when it is remembered that science is merely the 'art of the possible'[16] and does by no means aim at discovering the absolute 'truth' (which is at best an asymptotic concept not accessible via 'finite efforts'). This is not to deny the need for time evolution of even the most fundamental concepts of science (did not relativity and quantum theory eventually replace the classical Euclid–Newton concepts?), but the time scale for such evolutions is the result of a very complex feedback within the scientific community and cannot be determined on the basis of any insufficient efforts, howsoever reliable, from a few individual workers without a thorough debate within the community on their wider implications.

The first example concerns the famous expedition of Eddington in 1919 to verify Einstein's general relativity. Even a cursory perusal of the Collins–Pinch account would convince anyone that the 'raw data' were hardly discriminatory as between the Newton and the Einstein predictions (which differed by half). Yet a conscious value judgement had to be introduced by Eddington (with active support from the Astronomer–Royal) to decide in favour of the Einstein (not Newton) prediction. At the hands of a lesser mortal such a 'doctoring' of data would either be ignored or (if, the worker has a good reputation) provoke violent protests. Indeed, in the present age such 'poor' data would simply be 'unacceptable' in the first place, a recent example being the 'monopole' event of Cabrera[17]. A more substantial evidence of the demands of modern rigour in the processing of data is provided by the 'processing time' taken by the CDF team work before finally deciding to publish their results[13], even though the crucial nature of their findings (for which the entire community had been waiting with bated breath) would have ensured a ready acceptance. A greater boldness (and daring!) had been shown by Carlo Rubbia a decade earlier, when, as leader of the UA2 team, he had decided to publish their findings somewhat in advance of the final results, for which the entire high-energy community had been waiting anxiously.

However, Eddington's expedition was in the Rutherford era when science and its practitioners used to be regarded with far greater awe and reverence than is the case today. And, with his formidable standing within the scientific community, he was able to push through his 'interpretation' successfully, as the circumstantial evidence was overwhelmingly in its favour, though not necessarily on the basis of the actual data obtained from the expedition. The important lesson from this episode is that an experimental 'confirmation' of a theory in the standard orthodox fashion – the bedrock of 'objectivity' in science – is not always relevant for 'belief in the theory'! In retrospect, Eddington turned out to be right, but for a 'wrong' reason, since it would not have mattered much what the data showed. On the other hand, his deep insights in physics, howsoever buttressed by his high prestige, had their own limitations as were tellingly confirmed by his own student Chandrasekhar's subsequent 'victory' over the Professor's (superior?) judgement on the now famous *Chandrasekhar limit*. And yet the damage that the Professor's initial verdict had caused to his student's career in the interim period is a matter for history[18]. Lesson? Personal authority as an instrument of subjectivity must not be pushed too far, since its credibility is not unlimited! Failure of 'institutional authority' is already illustrated (see above) by the failure of proton decay within the time limits set by the grand unification theory.

Yet subjectivity has been known to 'pay' if combined with a right 'dose' of intuition together with establishment's support, in both of which the actual 'experimental' evidence, especially in the short run, often plays a marginal role. Thus, in the famous 'experimental' debate between Pouchet and Pasteur[15] on the origin of life, Pasteur was against the spontaneous generation theory advocated by Pouchet. Although Pouchet's experiments seemed to support his theory, Pasteur would invariably come up with some 'catch' in the experimental material and/or set-up. He even 'discarded' some of his own experiments which showed a bias for 'spontaneous generation', relying more on his intuitive perceptions than on the actual observations. And he was supported by the Paris-based Academie des Sciences – the establishment – even though the latter's reasons were anything but objective. That Pasteur was upheld by later, more precise, results is an entirely different matter.

A similar incident of more recent origin concerns Fairbank's (an experimental physicist of considerable reputation) 'discovery of the quark'[19], a result which was hotly disputed by Morpurgo, who was so sure of his ground that he even ventured to suggest that it was an artefact of the 'mode of observation'[20]! Finally, at the Rochester Conference of 1960, during the heydays of the V–A theory, Feynmann simply refused to believe that the theory could be wrong, even though some helium-6 data were 'strongly supporting' the scalar-tensor theory. Such examples can be multiplied.

All these examples illustrate the more 'noble' aspects of subjectivity that relate to a 'value judgement' in which deep intuition plays a vital role.

Conclusion

In the foregoing we have tried to depict several facets of subjectivity, from value judgement to human interactions, that are employed in the actual practice of science. While the traditional picture of science is one of 'strictly objective' interplay of theory and experiment within the rules of the game, subjectivity enters in a vital way as a sort of value judgement in the worker's own perceptions of the same. This 'abstract' aspect of subjectivity is an intensely personal perception which remains hidden from view but may be reflected in his choice of problems and the tenor of his writings. The more visible aspect of subjectivity, on the other hand, enters through human interaction, wherein science is a part of the community's larger concerns, and hence is subject to all standard types of socioeconomic and sociopolitical pressures. Sandwiched between these dominantly 'subjective' elements, the 'objective' part of his practice of science is more or less a routine affair, appearing essentially as an end product of his investigation, while the vital decisions are controlled by the former.

1. Holton, G., *From Kepler to Einstein*, Harvard University Press, Harvard, 1973.
2. Quoted in ref. 1.
3. Mitra, A. N., *Curr. Sci.*, 1993, **64**, 162.
4. Mitra, A. N., *Phys. News*, 1990, **21**, 7.
5. Mitra, A. N., *Pramana – J. Phys.*, 1986, **27**, 73.
6. Gell-Mann, M., *Quark and the Jaguar*, Little, Brown and Co., London, 1994.
7. Weisskopf, V. F., *Phys. Today*, 1970, **23**, 17.
8. Kuhn, T., *The Structure of Scientific Revolutions*, University of Chicago Press, Chicago, 1973.
9. Davies, P., in *Doing Science* (ed. Brockman, J.), Prentice Hall, New York, 1988–1991.
10. Omnes, R., *Rev. Mod. Phys.*, 1992, **64**, 339.
11. Notalle, L., *Fractal Space-Time and Microphysics*, World Scientific, 1993.
12. Morris, R., in ref. 9.
13. CDF Collaboration, *Phys. Rev. Lett.*, 1994, **72**, 1977.
14. Mitra, A. N. and Ramanathan, R., *Z. Phys.*, 1984, **C22**, 351.
15. Collins, H. and Pinch, T., *The Golem*, Cambridge University Press, Cambridge, 1993.
16. van Hove, Leon, in a private conversation, 1970.
17. Cabrera, B. *et al.*, *Phys. Rev. Lett.*, 1983, **51**, 1933.
18. Wali, K. C., *S. Chandrasekhar's Biography*, Penguin, New Delhi, 1994.
19. Fairbank, W., *Phys. Rev. Lett.*, 1977, **38**, 1011.
20. Morpurgo, G., private communication, 1979.

A. N. Mitra is in the National Institute of Advanced Studies, IISc Campus, Bangalore 560 012, India.

Edited and published by Prof. S. Ramaseshan and Prof. P. Balaram, Current Science Association, Bangalore 560 080.
Typeset by R D Typesetters, Bangalore 560 010. Printed by Lotus Printers, Bangalore (Ph: 3380167).

We'll help you get there

If you're selling science books/journals or laboratory products, or wish to recruit scientists in your establishment, then

you're sure to benefit by advertising here.

Because *Current Science* reaches nearly every university, scientific institution and industrial R&D unit in India. What's more, it's read by hundreds of individual subscribers – students, doctoral scientists and professionals in virtually every field of scientific activity in India.

Current Science has the largest circulation among scientific journals in India.

And to give you more impact, we're now bigger, and better.

Write now, or send your copy to

CURRENT SCIENCE

C.V. Raman Avenue, P.B. No. 8001,
Bangalore 560 080.

NEW Display Advertisement Rates*

		Tariff (rupees)					
		Inside pages		Inside cover pages		Back cover page	
No. of insertions	Size	B&W	Colour	B&W	Colour	B&W	Colour
1	Full page	5,000	10,000	7,000	12,000	10,000	15,000
	Half page	3,000	5,000	–	–	–	–
6	Full page	25,000	50,000	35,000	60,000	50,000	75,000
	Half page	15,000	25,000	–	–	–	–
12	Full page	50,000	1,00,000	70,000	1,20,000	1,00,000	1,50,000
	Half page	30,000	50,000	–	–	–	–
24	Full page	90,000	1,80,000	–	–	–	–
	Half page	50,000	90,000	–	–	–	–

SEND PAYMENT BY BANK DRAFT WITH ORDER

*Effective from 1 January 1995

CURRENT SCIENCE

A fortnightly journal of research

Editors: S. Ramaseshan and P. Balaram

Editorial Board

S. Arunachalam, Central Electrochemical Research Institute, Karaikudi 623 006
Asis Datta, School of Life Sciences, Jawaharlal Nehru University, New Delhi 110 067
D. Balasubramanian, Centre for Cellular and Molecular Biology, Hyderabad 500 007
J. Chandrasekhar, Department of Organic Chemistry, Indian Institute of Science, Bangalore 560 012
R. Gadagkar, Centre for Ecological Sciences, Indian Institute of Science, Bangalore 560 012
K. N. Ganeshaiah, Department of Plant Genetics and Breeding, University of Agricultural Sciences, Bangalore 560 065
R. V. Hosur, Tata Institute of Fundamental Research, Bombay 400 005
N. V. Joshi, Centre for Ecological Sciences, Indian Institute of Science, Bangalore 560 012
Jyotirmoy Das, Indian Institute of Chemical Biology, Calcutta 700 032
C. C. Kartha, Division of Cellular and Molecular Cardiology, Sree Chitra Tirunal Institute for Medical Sciences and Technology, Thiruvananthapuram 695 011
S. S. Krishnamurthy, Inorganic and Physical Chemistry Department, Indian Institute of Science, Bangalore 560 012
S. C. Lakhotia, Department of Zoology, Banaras Hindu University, Varanasi 221 005
K. Muralidhar, Department of Zoology, Delhi University, Delhi 110 007
R. Narasimhan, CMC Limited, Bangalore 560 001
K. Neelakantan, ANURAG (Advanced Numerical Research and Analysis Group), Hyderabad 500 258
R. Nityananda, Raman Research Institute, Bangalore 560 080
T. Padmanabhan, Inter-University Centre for Astronomy and Astrophysics, Pune 411 007
B. R. Pai, Propulsion Division, National Aerospace Laboratories, Bangalore 560 017
T. J. Pandian, School of Biological Sciences, Madurai Kamaraj University, Madurai 625 021
Y. S. Rajan, Department of Science and Technology, New Delhi 110 016
M. Ramakrishnan, Geological Survey of India, Calcutta 700 016
A. V. Ramani, T. T. K. Pharma Ltd., Bangalore 560 025
S. Ramaswamy, Department of Physics, Indian Institute of Science, Bangalore 560 012
A. K. Raychaudhuri, Department of Physics, Indian Institute of Science, Bangalore 560 012
S. R. Shetye, Physical Oceanography Division, National Institute of Oceanography, Dona Paula, Goa 403 004
V. Siddhartha, 51 Bharati Nagar, New Delhi 110 003
V. Srinivas, School of Mathematics, Tata Institute of Fundamental Research, Bombay 400 005
G. Srinivasan, Raman Research Institute, Bangalore 560 080
R. Uma Shaanker, Department of Crop Physiology, University of Agricultural Sciences, Bangalore 560 065
K. S. Valdiya, Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore 560 064
G. Venkataraman, Department of Physics, Sri Sathya Sai Institute of Higher Learning, Prasanthi Nilayam 515 134
M. Vidyasagar, Centre for Artificial Intelligence and Robotics, Bangalore 560 001
K. VijayRaghavan, TIFR Centre, Indian Institute of Science Campus, Bangalore 560 012

Current Science, founded in 1932, is published by the Current Science Association in collaboration with the Indian Academy of Sciences. The journal is also intended as a medium for communication and discussion of important issues that concern science and scientific activity. All articles published in *Current Science*, especially editorials, opinions and commentaries, letters and book reviews, are deemed to reflect the individual views of the authors and not the official points of view, either of the Current Science Association or of the Indian Academy of Sciences.

Working Committee, Current Science Association

P. Balaram, Indian Institute of Science, Bangalore
R. Chidambaram, Department of Atomic Energy, Bombay
V. K. Gaur, CSIR Centre for Mathematical Modelling and Computer Simulation, Bangalore
K. Kasturirangan, ISRO Satellite Centre, Bangalore
P. K. Kaw, Institute for Plasma Research, Gandhinagar
N. Kumar, Raman Research Institute, Bangalore (*Vice-President*)
N. V. Madhusudana, Raman Research Institute, Bangalore (*Secretary*)
A. Mani, Field Research Unit, Indian Institute of Tropical Meteorology, Bangalore (*Past-President*)
R. A. Mashelkar, Council of Scientific and Industrial Research, New Delhi
R. Narasimha, Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore
G. Padmanaban, Indian Institute of Science, Bangalore (*Vice-President*)
R. Ramachandran, The Institute of Mathematical Sciences, Madras
P. Rama Rao, Defence Research and Development Organization, New Delhi
S. Ramaseshan, Raman Research Institute, Bangalore (*Treasurer*)
M. S. Valiathan, Manipal Academy of Higher Education, Manipal (*President*)

Editorial Office

Editorial staff

Chandrika Ramesh, G. Madhavan, T. D. Mahabaleswara, C. S. Ravi Kumar, M. S. Venugopal

Circulation and Accounts

Peter Jayaraj, Ranjini Mohan, B. Sethumani, Shanthi Bhaskar, B. K. Shivaramiah

Annual subscription 1995

	Institutional	Personal
India	Rs 350	Rs 100
All other countries (AIR MAIL)	US $175	US $50

Single copies other than special issues: Rs 35/US $15

Editorial Office: CURRENT SCIENCE
C. V. Raman Avenue
P.B. 8001
Bangalore 560 080, India
Telephone: 91-80-3342310
Telex 845-2178 ACAD IN
Fax: 91-80-3346094

Copyright © 1995, Indian Academy of Sciences

INFORMATION FOR CONTRIBUTORS

GENERAL

All manuscripts should be addressed to the Editor, *Current Science*, P. B. No. 8001, C. V. Raman Avenue, Bangalore 560 080. Submission of an article will be held to imply that it has not been previously published and is not under consideration for publication elsewhere; and further, that if accepted, it will not be published elsewhere. *Three copies of contributions of all categories* are required, with a letter of transmittal giving (i) names and complete addresses of the authors and (ii) title of the contribution and the category in which it is submitted (see below).

Current Science is a multidisciplinary journal and therefore research and review papers of general significance that are written clearly and well organized will be given preference. All papers will be first assessed by a Reviewing Editor. Papers found unsuitable in terms of the overall requirements of the journal will be returned to the authors. The others will be sent for detailed review. *Both solicited and unsolicited material will be reviewed.* Authors of these papers will be notified of acceptance, rejection, or need for revision of the paper. Returned papers cannot be resubmitted.

Illustrations and other material to be reproduced from other publications must be properly credited; it is the authors' responsibility to obtain permission for reproduction (copies of letters of permission should be sent).

CATEGORIES OF MANUSCRIPT

General articles (not exceeding 5000 words) discuss current trends in research in a field that will be of interest to readers outside the field; interdisciplinary topics: science policy and administration; or some aspect of the application of science and technology to human needs or the impact of science and technology on society/ecosystems/life. They should include a summary not exceeding 100 words, introductory paragraph(s), brief subheads at appropriate places to point to what follows, illustrations that will help a general reader, and references.

Review articles (not exceeding 5000 words) are expected to survey and discuss recent developments in a field. They should be well focused and organized, and avoid a general, 'textbook' style.

Research articles (not exceeding 4000 words) should report research results of fairly major significance. They should include an abstract not exceeding 100 words, introductory paragraph(s), and brief subheads.

Research communications (not exceeding 2000 words) should contain important findings that are novel and of fairly broad interest. They should include a brief abstract and an introductory paragraph. Text should not be broken up under subheads.

Correspondence includes letters, not exceeding 500 words, that are of general interest to scientists. All letters cannot be published.

Scientific correspondence contains technical comments, including those on articles or communications published in *Current Science* within the previous six months. Letters may be reviewed and edited.

Research news articles are intended to inform nonspecialists about recently published advances or important findings discussed at a meeting. Authors should also send a copy of the paper(s) on which the article is based. Meeting reports should avoid merely listing brief accounts of topics discussed, and must convey to readers the significance of an important advance.

Research accounts articles are intended to be personalized reviews of research from the authors' own laboratory, based on a body of published work. The articles must provide appropriate background to the area in a concise introduction, which should also serve to place the author's work in proper perspective. Articles will normally not exceed 8 to 10 printed pages.

Opinion articles present views on issues related to science and scientific activity. **Commentary** articles should contain expository notes on issues related to science and scientific activity.

Book reviews. Unsolicited reviews will also be considered. Reviews that merely 'list' brief descriptions of the contents cannot be published. Reviews should have 'context' and convey some information about the subject of the book.

Historical commentary and notes inform readers about interesting aspects of personalities or institutions of science or about watershed events in the history/development of science; most are expected to relate to India. Illustrations are welcome. Brief items will also be considered.

MANUSCRIPT PREPARATION

Manuscripts should be typed double-spaced on one side of white bond paper (21×28 cm). The pages should be numbered consecutively, starting with the title page and through the text, reference list, tables and figure legends. The **title** should be brief, specific and amenable to indexing. Not more than five **keywords** should be indicated separately; these should be chosen carefully and must not be phrases of several words. **Summary** and **abstract** should not have more than 100 words and should convey the main point of the paper, outline the results and conclusions, and explain the significance of the results.

Text. All papers should have a brief introduction. The text should be intelligible to readers in different disciplines and technical terms should be defined. Tables and figures should be referred to in numerical order. All **symbols** and **abbreviations** must be defined, and used only when absolutely necessary. Superscripts and subscripts and ambiguous characters should be clearly indicated. **Units of measure** should be metric or, preferably, SI. Methods should, as far as possible, be described briefly in appropriate table and figure legends.

Figures. In the case of line drawings one set of originals (without any lettering) is sufficient, with two copies containing lettering. In the case of photographs good prints are required with each copy of the manuscript; photocopies are not acceptable. Line drawings should be roughly twice the final printed size. The correct orientation should be indicated if not clear.

Photomicrographs and other photographs that require it must have a scale bar, which should be defined clearly in the legend. Primary data should be submitted as far as possible (e.g. actual photographs of electrophoretic gels rather than idealized diagrams).

References should be numbered in the order in which they appear, first through the text and then through table and figure legends. The following are examples of ways of writing references.

1. Mukundan, T. and Kishore, K., *Curr. Sci.*, 1991, **60**, 355–362.
2. Constantine, G., in *Biology of Bats* (ed. Wimsatt, W. A.), Academic Press, New York, 1970, vol. 1, pp. 319–322.

Acknowledgements should be brief. Footnotes are not allowed except to identify the corresponding author if not the first.

Cover photographs. Good photographs (colour or black and white) that pertain to a submitted paper will be considered for use on the cover. Good *prints* and a legend should be submitted with the manuscript. In the case of a colour picture, a transparency will be required for printing if accepted.

PROOFS AND PUBLICATION

Two sets of galley proofs are sent to the corresponding author. A reprint order form accompanies the proofs.

Founded: 1932

ISSN: 0011-3891

CURRENT SCIENCE

Volume 69 Number 10

25 November 1995

791 **In this issue**

NEWS

793 **Pioneers of ozone study bag chemistry Nobel**
S. Parthiban

797 **The 1995 Nobel Prize in Medicine: A tribute to the power of formal genetics**
S. C. Lakhotia

CORRESPONDENCE

801 **If education is important, so must be the teachers!**
Lambodar P. Singh

COMMENTARY

802 **India's scientific community: A cultural crisis?**
A. N. Mitra

SCIENTIFIC CORRESPONDENCE

809 **Ethephon enhances karaya gum yield and wound healing response: A preliminary report**
M. N. B. Nair, K. R. Shivanna and H. Y. Mohan Ram

810 **Occurrence of *Botryococcus* from the Lower Cretaceous rocks of Cauvery Basin, Tamil Nadu**
R. S. Patil, M. Rama Rao, M. N. Gururaja, P. Kumaraguru and B. Kanishkan

811 **Enechelon faults along West Coast of India and their geological significance**
Sm. Ramasamy

814 **Interpretation of nucleotide protein sequence data: Some pitfalls**
Chitra Dutta and Jyotirmoy Das

GENERAL ARTICLE

816 **Fifty years of the exact solution of the two-dimensional Ising model by Onsager**
Somendra M. Bhattacharjee and Avinash Khare

SPECIAL SECTION: AIDS IN ASIA

822 **Introduction**
David E. Bloom

823 **Biomedical aspects of HIV and AIDS**
John M. Dwyer

828 **Estimation and projection of HIV infection and AIDS cases**
James Chin

833 **The problem of HIV and AIDS in India**
Shiv Lal and B. B. Thakur

835 **The HIV and AIDS epidemic in the Western Pacific**
Rabin M. Sarda and George Petersen

840 **Heterosexual risk behaviour in Asia: The implications for HIV/AIDS**
Tim Brown and Werasit Sittitrai

849 **The economics of the AIDS epidemic: Lessons from the United States**
Sherry Glied

857 **Making the blood supply safe from HIV: The Sri Lankan experience**
Nandrani S. De Zoysa

861 **Ethical issues posed by HIV testing and screening**
Ronald Bayer

865 **The AIDS epidemic and economic policy analysis**
David E. Bloom and Ajay Mahal

870 **Contributors to the Special Section**

RESEARCH COMMUNICATIONS

871 **Standard potential of hydrogen electrode**
S. Parthasarathy, K. Ramya and V. K. Venkatesan

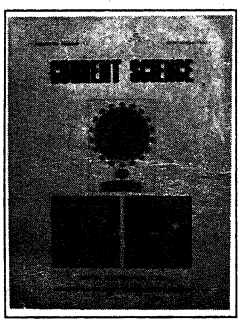
874 **Selective denaturation of cytochrome c oxidase by ionic surfactants: Depletion of both the heme *a* residues from the enzyme**
Tapan Kanti Das and Shyamalava Mazumdar

877 **Evaluation of haemorheological, biochemical and microcirculatory parameters in young smokers**
Ranjeet S. Ajmani, R. R. Puniyani and Varjeet S. Khanuja

BOOK REVIEWS

881 **Law, Ethics and HIV,** *reviewed by* N. R. Madhava Menon

882 **The Collected Papers of Albert Einstein – Volume 5 – The Swiss Years: Correspondence, 1902–1914,** *reviewed by* N. Mukunda



COVER. *Top*: A diagrammatic representation of the human immunodeficiency virus (HIV). The central core of the virus contains the genetic information that the virus needs to reproduce itself. The highly complex egg-shaped projections covering the surface of the virus allow it to attack CD4 lymphocytes in a unique fashion. *Below left*: The retina of a patient with AIDS: the retina is usually pink and blood vessels can be seen running over its surface. The white patches represent infection with cytomegalovirus. *Below right*: The eye of the same patient a week later, showing how rapidly the infection can spread when T cells are unavailable for defence. See Special Section. [Reproduced with permission from *The Body at War* by John M. Dwyer, Publishers Allen and Unwin.]

Indexed in CURRENT CONTENTS/GEOBASE/CHEMICAL ABSTRACTS

The editors thank Prof. David E. Bloom, Harvard Institute for International Development, 1 Eliot Street, Cambridge, MA 01238, USA, for agreeing to be Guest Editor of the Special Section.

Single copy of this issue, Rs 150/US $20.

In this issue

AIDS

Almost fifteen years have passed since the disease AIDS surfaced in the United States. This immune deficiency state which rendered patients susceptible to opportunistic infections was soon recognized to be a sexually transmitted disease, with the causative agent being a retrovirus. Originally thought to be limited to homosexual populations, AIDS was soon to assume the dimensions of a major public health problem when heterosexual transmission and infection via blood transfusions and contaminated syringes in drug addicts were established. There is no satisfactory treatment for AIDS and approaches to limiting the spread of the human immunodeficiency virus (HIV) are still to take hold. The magnitude of the AIDS problem in India appears to be relatively small at present; malaria, tuberculosis and diarrhoeal diseases may still be more widespread. However, the dangers of underestimating an insidious enemy are all too clear. A special section in this issue, edited by David Bloom highlights the complexity of the problems involved in addressing the growing AIDS epidemic.

P. Balaram

Fifty years after a landmark

The melting of ice and the loss of magnetism by iron on heating are two examples of what physicists call phase transitions. In the second example, the order becomes zero continuously at a critical temperature. Various approximate models for this behaviour have been proposed, starting with van der Waals in the last century. On close examination, they proved inconsistent, both with the best experiments and internally. Fifty years ago, Lars Onsager found an exact mathematical solution of one model of a phase transition — a brilliant discovery which changed the course of the subject. This historic event, the prelude to it and its aftermath are recounted for our readers by S. M. Bhattacharjee and A. Khare (**page 816**).

R. Nityananda

We'll help you get there

If you're selling science books/journals or laboratory products, or wish to recruit scientists in your establishment, then

you're sure to benefit by advertising here.

Because *Current Science* reaches nearly every university, scientific institution and industrial R&D unit in India. What's more, it's read by hundreds of individual subscribers – students, doctoral scientists and professionals in virtually every field of scientific activity in India.

Current Science has the largest circulation among scientific journals in India.

And to give you more impact, we're now bigger, and better.

Write now, or send your copy to

CURRENT SCIENCE

C.V. Raman Avenue, P.B. No. 8001,
Bangalore 560 080.

NEW Display Advertisement Rates*

		Tariff (rupees)					
		Inside pages		Inside cover pages		Back cover page	
No. of insertions	Size	B&W	Colour	B&W	Colour	B&W	Colour
1	Full page	5,000	10,000	7,000	12,000	10,000	15,000
	Half page	3,000	5,000	–	–	–	–
6	Full page	25,000	50,000	35,000	60,000	50,000	75,000
	Half page	15,000	25,000	–	–	–	–
12	Full page	50,000	1,00,000	70,000	1,20,000	1,00,000	1,50,000
	Half page	30,000	50,000	–	–	–	–
24	Full page	90,000	1,80,000	–	–	–	–
	Half page	50,000	90,000	–	–	–	–

SEND PAYMENT BY BANK DRAFT WITH ORDER

*Effective from 1 January 1995

CURRENT SCIENCE

Volume 69 Number 10 25 November 1995

NEWS

Pioneers of ozone study bag chemistry Nobel

The 1995 Nobel prize in chemistry has been awarded jointly to Paul J. Crutzen, Department of Atmospheric Chemistry and Biogeochemistry, Max-Planck-Institute for Chemistry, Mainz, Germany; Mario J. Molina, Department of Earth, Atmospheric and Planetary Sciences and Department of Chemistry at Massachusetts Institute of Technology, USA; and F. Sherwood Rowland, Department of Chemistry at the University of California, Irvine, USA.

The laureates are recognized for their pioneering work on the subject of formation and decomposition of ozone. Crutzen was instrumental in establishing the importance of nitrogen oxide chemistry. Rowland and Molina were the first to warn that chlorofluorocarbons (CFCs) released into the atmosphere were depleting the ozone layer. By explaining the chemical mechanisms that threaten to destroy the protective ozone screen surrounding the earth, the three scientists have contributed to our salvation from a global environmental damage that could have serious ill-effects.

The ozone layer[1-3]

Ozone, a three-atom molecule of oxygen, occurs naturally as a gas in trace amounts in the Earth's atmosphere. It is found primarily in a layer between about 12 and 50 km above the Earth's surface, in the region of the atmosphere called the stratosphere. The ozone layer is thin, fragile and almost invisible. Yet it plays an exceptionally fundamental role in determining the habitability of the Earth.

A one per cent decrease in ozone allows a two per cent increase in UV radiation that reaches the Earth. Effects from increased UV radiation include an increased risk of skin cancer, cataracts and immune deficiencies in humans, as well as damage to crops and the plankton that supports marine life. While the stratospheric ozone plays a beneficial role, the surface-level tropospheric ozone is one of the most harmful pollutants, damaging plants and humans. The tropospheric ozone is produced through smog type of reactions[4], a familiar problem in the atmosphere of many cities around the world.

The first qualitative photochemical theory for the formation and decomposition of ozone in the atmosphere was formulated by Sidney Chapman[5] in 1930. According to this theory, ozone is produced in the middle and upper atmosphere through dissociation of molecular oxygen by sunlight; the free oxygen atoms combine with other O_2 molecules to form ozone.

$$O_2 + UV(\lambda < 242\ nm) \rightarrow O + O$$
$$O + O_2 + M \rightarrow O_3 + M,$$

where M is a third body required to carry away the energy released in the combination reaction. The triple collision is rare at very high levels (above 60 km), because there are so few molecules. Thus, there is almost no ozone above 60 km. On the other hand, there is almost no atomic oxygen, O, below approximately 16 km, so the formation of ozone is unlikely there, too. These are the reasons most of the ozone is found in the stratosphere.

Early warnings

Chapman's theory, however, showed appreciable deviations from measurements. The calculated ozone contents were considerably higher than the observed ones. To account for this, Bates and Nicolet[6] drew attention to the role played by H, OH and HO_2 (products of UV photolysis of water vapour) in the catalytic reduction of odd oxygen above 60 km. But the calculations again showed very large concentrations of ozone than the observations.

Paul Crutzen[7] in 1970 suggested that additional important processes must be taken into account in order to correctly describe the photochemistry of the atmospheric ozone. Crutzen uncovered an important aspect that nitrogen chemistry accounts for a significant fraction of ozone loss in the natural atmosphere, particularly above 25 km.

$$NO + O_3 \rightarrow NO_2 + O_2$$
$$NO_2 + O \rightarrow NO + O_2$$
$$\text{net: } O_3 + O \rightarrow 2O_2.$$

Using available satellite data, Crutzen provided a much more quantitative



Crutzen was born in 1933 in Amsterdam. He obtained his doctorate in Meteorology in 1973 from Stockholm University. He is a member of the Royal Swedish Academy of Sciences, the Royal Swedish Academy of Engineering Sciences and Academia Europaea.

understanding of the sources of NO_x via N_2O oxidation

$$N_2O + O \rightarrow 2\ NO.$$

The primary sources of N_2O include oceans, combustion of fossil fuels, fertilized agriculture fields, tropical and subtropical forests and woodlands. N_2O diffuses upwards from troposphere to the stratosphere to give NO which in turn reacts with ozone. The connection demonstrated by Crutzen[8] between tropical biomass burning and the ozone concentration is one of the motives for the recent rapid development of research on global biogeochemical cycles.

Johnston[9] rediscovered the NO_x–O_3 sequence and made a connection to supersonic transport emissions. Supersonic civil aircraft can fly across the ocean at supersonic speeds, but cannot fly across land because of the sonic boom noise. These aircrafts would be capable of releasing nitrogen oxides right in the middle of the ozone layer at the altitudes of 20 km. Crutzen's and Johnston's work gave rise to a very intensive debate among researchers as well as among technologists and decision-makers. Since then, several multidimensional models[10] have been developed to assess the potential environmental effects of aircraft operations in the stratosphere. But it must be pointed out that opposition to the 1970s supersonic fleet came primarily because of the very high development cost and not just the NO_x engine emissions.

The CFC affair

The next vital step towards a deeper understanding of ozone chemistry was in 1974, when Rowland[11] and his then postdoc, Molina, established the possibility of major stratospheric ozone depletion from CFCs. During the early seventies, CFCs were used as tracers in the troposphere because of their lack of reactivity. According to 'Rowland–Molina hypothesis' as it is now often described, the release of atomic chlorine from the solar UV photolysis of CFCs triggers catalytic chain reaction that causes significant depletion of stratospheric ozone layer.

CFCs were invented in 1928 as safe alternatives to ammonia and sulphur dioxide refrigerants. Since then the uses of CFCs have been expanded to include aerosol propellants, air conditioning, refrigerants, foam-blowing agents, and as solvents in electronic industries[12]. They are volatile, nonflammable, noncorrosive, nonexplosive, low toxic and chemically inert, and hence safe to work with. The very lack of chemical reactivity that makes CFCs commercially useful also allows them to persist for many decades in the Earth's atmosphere. The most important CFCs[13] for ozone depletion are: trichlorofluoromethane - $CFCl_3$ (CFC-11), dichlorodifluoromethane - CF_2Cl_2 (CFC-12), 1, 1, 2 trichlorotrifluoroethane - $CF_2ClCFCl_2$ (CFC-113) and carbon tetrachloride (CCl_4).

Rowland and Molina realized that the chemically inert CFCs could be

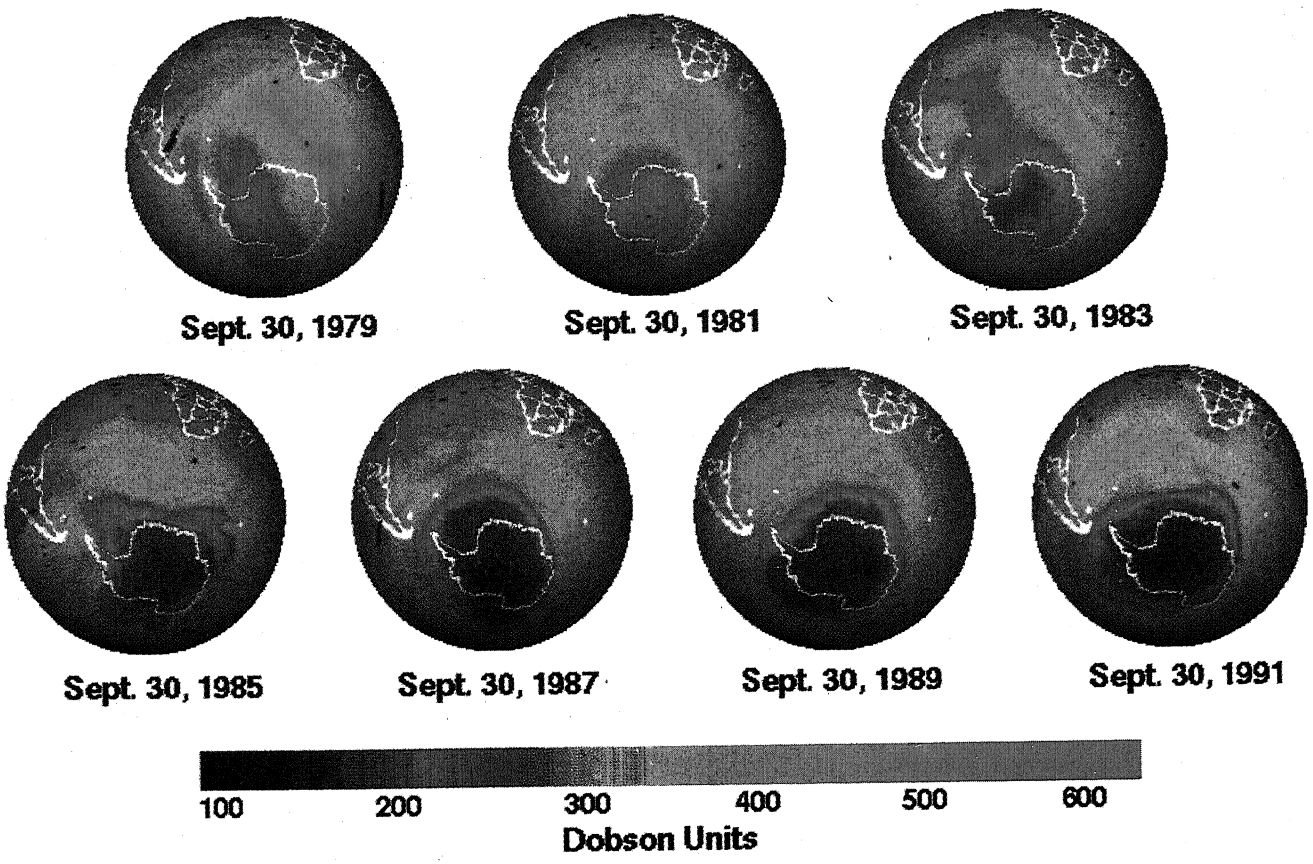


Figure 1. Progressive thinning of ozone layer as revealed by TOMS data.

transported to the stratosphere and decomposed by solar UV radiation.

$$CCl_2F_2 + UV(\lambda < 220 \text{ nm}) \rightarrow Cl + CClF_2.$$

After they are photolysed, most of the chlorine eventually ends up as hydrogen chloride, HCl, or chlorine nitrate, $ClONO_2$. These are called reservoir species. Alternatively, the chlorine atoms released could participate in the catalytic removal of ozone. The simplest mechanism suggested by Rowland and Molina is

$$Cl + O_3 \rightarrow ClO + O_2$$
$$ClO + O \rightarrow Cl + O_2$$
$$\text{net: } O + O_3 \rightarrow 2O_2.$$

Note that the role of chlorine atoms is that of a catalyst. Thus, an atom, once released, can participate in the O_3 removal as long as it stays in the atmosphere. The process is even more dramatic for bromine – it has no stable reservoirs; so the Br atom is always available to destroy ozone. On a per-atom basis, Br is 10–100 times as destructive as Cl. In contrast, fluorine is innocuous in the stratosphere. Most of it ends up as its reservoir – hydrogen fluoride (HF), which is extremely stable. Since chlorine and bromine radicals both cause ozone destruction, it comes as no surprise that iodine can also do so.

Hole in the ozone shield

Antarctica is the southernmost continent of the world, and literally speaking, one of the most remote places on the Earth. But here occurred the first large-scale ozone loss. Public concern was aroused in 1985 when atmospheric scientists of the British Antarctic Survey[14] announced their startling discovery of an 'ozone hole' over Halley Bay, Antarctica. Other scientists confirmed the discovery and reported that the region of ozone depletion was actually wider than the continent. These findings led to signing of a treaty in September 1987 on the protection of the ozone layer, known as Montreal Protocol, the first global environment protection agreement.

The quantitative aspects of the ozone hole, however, could not be explained by dynamical processes or by gas phase chemical reactions. In fact, atmospheric models employing the gas-phase chemistry estimated only a 1–2% loss of ozone, significantly less than what has been observed. Such findings prompted the ozone researchers to suspect that an alternative mechanism must exist that could accelerate the decomposition of ozone. As it turned out, Rowland[15] along with Solomon, Garcia and Wuebbles identified this mechanism as chemical reactions on the surface of cloud particles in the stratosphere. The clouds convert two so-called reservoir species of chlorine, HCl and $ClONO_2$, into more reactive chlorine (Cl_2) and hypochlorous acid (HOCl):

$$HCl + ClONO_2 \rightarrow HNO_3 + Cl_2$$
$$H_2O + ClONO_2 \rightarrow HNO_3 + HOCl.$$

The nitric acid that is produced sticks to cloud particles whereas the Cl_2 and HOCl fly off into the gas phase. Further, Crutzen[16] and Arnold emphasized that the clouds absorb nitrogen oxides and thus it allows the abnormally high chlorine to persist.



Rowland was born in Delaware, USA in 1927. He earned his B S degree from Ohio Wesleyan University and his M S degree from University of Chicago. In 1952, Rowland received his Ph D from the University of Chicago and taught at Princeton University and the University of Kansas before going to Irvine in 1964 as founding Chairman of its chemistry department. Before the publication of the Rowland–Molina papers on CFC problem, his research centered on the chemical reactions of radioactive isotopes. Currently, Rowland is Emeritus Professor in the University of California and he is also Foreign Secretary of the US National Academy of Sciences.

$$NO_2 + NO_3 + M \rightarrow N_2O_5 + M$$
$$N_2O_5 + H_2O \rightarrow 2HNO_3.$$

That is, the conversion blocks the following reaction

$$ClO + NO_2 \rightarrow ClONO_2.$$

As a consequence of these studies, an exciting new branch of atmospheric chemistry has emerged: 'heterogeneous' chemical reactions on particle surfaces.

In the stratosphere, a stream of air known as polar vortex circles Antarctica in winter. Air trapped within this vortex becomes extremely cold during the polar night. Temperatures drop low enough to form clouds even in the very dry stratosphere. Particles of nitric acid trihydrate (type I polar stratospheric clouds (PSCs)) first condense at about 193 K; water ice particles (type II clouds) form when the temperature falls still lower, to about 187 K. These clouds freeze-out water vapour (dehumidification) and remove nitrogen oxides (de-nitrification) by adsorption of HNO_3 into the PSCs[17]. As mentioned earlier, the PSCs provide a reaction surface for transferring inactive chlorine to an active stable form ready to react once sunlight appears in the spring. The arrival of UV radiation photolyses the Cl_2 to radicals that can catalyse ozone destruction.

Several proposals have been made for the catalytic mechanism linking halogen radicals to the chemical destruction of ozone[18]. For example, McElory[19] and coworkers suggested ClO–BrO mechanism while Solomon[15] and colleagues proposed ClO–HO_2 sequence in addition to the classical catalytic couplet originally introduced by Molina and Rowland. Nevertheless, a chlorine chain[20] suggested in 1987 by Mario Molina and his wife Louisa J. Molina, involving ClO dimer formation at low temperatures followed by photolysis and thermal decomposition is now thought to account for the massive ozone destruction. The proposed mechanism follows.

$$Cl_2 + h\nu \rightarrow 2Cl$$
$$Cl + O_3 \rightarrow ClO + O_2$$
$$Cl + O_3 \rightarrow ClO + O_2$$
$$ClO + ClO + M \rightarrow ClOOCl + M$$
$$ClOOCl + h\nu \rightarrow Cl + ClOO$$
$$ClOO + M \rightarrow Cl + O_2 + M$$
$$\text{net: } 2O_3 \rightarrow 3O_2.$$

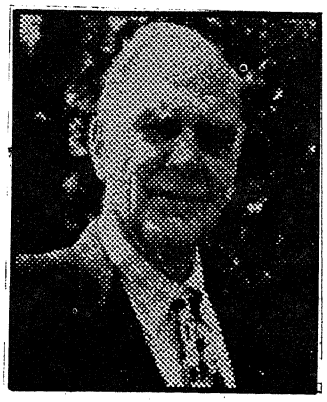
Molina[21] and coworkers also proposed a variation on this sequence in which the

photolysis of ClOOCl was replaced by thermal decomposition.

Basically, the overall effect of all these processes is to pre-condition the Antarctic stratosphere so that when sunlight returns to the stratosphere in spring two ozone molecules can be transformed into three molecules of oxygen through the above sequence of reactions. Once the spring is well-on, the Sun induces a dynamical breakdown of the stable vortex, and the 'chemical containment vessel', as it is sometimes called, is broken down and mixing of air with lower latitude occurs. PSCs evaporate and the atmosphere returns to a normal state.

Hence, in late winter, regions of ozone depletion around the pole begin to appear. Within a few weeks, the ozone hole is completely established. The hole disappears when the polar vortex finally breaks down after the spring Sun warms the air over Antarctica. Air then sweeps in from low latitudes, bringing nitrogen oxides that tie up the active chlorine and ozone that fills up the hole. The depletion first identified over Antarctica is now being observed to a lesser extent over the Arctic too, and is also much more evident at middle latitudes than has ever been predicted.

The Total Ozone Mapping Spectrometers (TOMS) aboard the NASA's Nimbus-7 spacecraft (launched on 24 October 1978) measures the total column ozone density of stratospheric ozone by observing back scattered solar UV radiation. The results of such measurements are usually presented in Dobson Units (1 DU = 2.69×10^{16} molecules of ozone cm^{-2}). Time series of TOMS images are shown in Figure 1, which illustrates the changes in total ozone over Antarctica at the beginning of every austral spring. The colour scale at the bottom shows the total ozone values in Dobson units. As a matter of fact, the roughly circular ozone hole rotates around the south pole at an offset, moving around the Antarctic continent and back to original position about once a week.



Mario Molina – the first Mexican to receive a scientific Nobel – was born in 1943 in Mexico City, Mexico. He received his Ph D in Physical Chemistry from the University of California, Berkeley. His wife Louisa is a Chinese and she is also a physical chemist. Mario is a member of the US National Academy of Sciences.

Skeptics, CFCs, and the ozone layer

When Rowland and Molina predicted in 1974 that chlorine from CFCs could chew away the ozone layer, some skeptics did not want to believe it at first. Further, the critics claimed that only small amounts of CFCs could enter the stratosphere because they are heavier than air (the molecular weight of CFC-11 (CCl_3F) is 137.5; that of air 29), adding that the chlorine is vented into the atmosphere by evaporation of sea water and eruptions of volcanoes.

Rowland made a broad attempt to confront the controversy and tackled the misconceptions concerning CFCs and ozone depletion. He argued that indoor experience with mass-dependent molecular diffusion is not applicable in the wind-driven atmosphere, in which the mixing process does not distinguish molecular masses[22]. If atmospheric gases segregate by weight, then the heavier molecules should dominate at lower altitudes. In reality, however, the relative proportions of Helium-4, Nitrogen-14, and Krypton-84 are unchanged from ground-level to about 100 km, despite the variation in molecular weight. Concerning the chlorine from sea water, it is in the form (NaCl) that is rapidly removed from the troposphere by rain. Volcanoes do eject chlorine as HCl during eruptions. But again, HCl is washed out by rain in the troposphere. Only rarely are eruptions so violent that they inject material directly into the stratosphere.

At last, the fierce debate is now over and the researchers are convinced that chlorine from CFCs is responsible for the ozone loss. Thanks to pioneering research by many researchers, among them Crutzen, Molina and Rowland, as well as Susan Solomon, James Anderson, Richard Stolarski, Ralph Cicerone and Donald Wuebbles.

Deservedly rewarded...

In recent years, Crutzen[23] has continued to make many other important contributions to the current understanding of atmospheric chemistry, both in the troposphere and stratosphere. Molina's[24] work at MIT (and University of California, Irvine and Jet Propulsion Laboratory before going to MIT) has mainly emphasized laboratory studies of heterogeneous reactions in relevance to atmospheric chemistry. Rowland's[25] work has primarily involved grab sample measurements of tropospheric gases, such as methane, hydrocarbons and halocarbons. Equally important, he urged scientists to reach beyond their subspecialities and attempt the difficult job of communicating sound science to the public.

Without a doubt, ozone researchers are all delighted and pleasantly surprised by the selection of environmental sciences for recognition by the Nobel committee. The Nobel prize to the trio of atmospheric scientists has elevated the whole area of ozone research. Indeed, it is the first Nobel prize for any environmentally related work. The Nobel committee, perhaps, is starting a trend to recognize scientists who have made fundamental contributions in applied sciences!

1. Rowland, F. S., *Annu. Rev. Phys. Chem.*, 1991, **42**, 731–768.
2. Jackman, C. H., *Environ. Sci. Technol.*, 1989, **23**, 1329–1332.
3. Wuebbles, D. J. and Kinnison, D. E., *Energy*, 1993, **18**, 1249–1262.
4. Thompson, A. M., *Science*, 1992, **256**, 1157–1165.
5. Chapman, S., *Mem. R. Meteorol. Soc.*, 1930, **3**, 103–125.
6. Bates, D. R. and Nicolet, M., *J. Geophys. Res.*, 1950, **55**, 301–327.
7. Crutzen, P. J., *J. Geophys. Res.*, 1971, **76**, 7311–7327.
8. Crutzen, P. J. and Andreae, M. O., *Science*, 1990, **250**, 1669–1678.
9. Johnston, H. S., *Science*, 1971, **73**, 517–522.

10. WMO, Scientific Assessment of Ozone Depletion: 1991, World Meteorological Organization Global Ozone Research and Monitoring Project-Report No. 25, Geneva, 1992.
11. Molina, M. J. and Rowland, F. S., *Nature*, 1974, **249**, 810–812.
12. McFarland, M., *Environ. Sci. Technol.*, 1989, **23**, 1203–1207.
13. Rowland, F. S., *Environ. Sci. Technol.*, 1991, **25**, 622–628.
14. Farman, J. C., Gardiner, B. G. and Shanklin, J. D., *Nature*, 1985, **315**, 207–210.
15. Solomon, S., Garcia, R. R., Rowland, F. S. and Wuebbles, D. J., *Nature*, 1986, **321**, 755–758.
16. Crutzen, P. J. and Arnold, F., *Nature*, 1986, **324**, 651–655.
17. Fahey, D. W., Kelly, K. K., Ferry, G. V., Poole, L. R., Wilson, J. C., Murphy, D. M., Lowenstein, M. and Chan, K. P., *J. Geophys. Res.*, 1989, **94**, 11299–11315.
18. Anderson, J. G., Brune, W. H., Lloyd, S. A., Toohey, D. W., Sander, S. P., Starr, W. L., Lowenstein, M. and Podolske, J. R., *J. Geophys. Res.*, 1989, **94**, 11480–11520.
19. McElory, M. B., Salawitch, R. J., Wofsy, S. F. and Logan, J. A., *Nature*, 1986, **321**, 759–762.
20. Molina, L. T. and Molina, M. J., *J. Phys. Chem.*, 1987, **91**, 433–436.
21. Molina, M. J., Tso, T. L., Molina, L. T. and Wang, F. C. Y., *Science*, 1987, **238**, 1253–1257.
22. Rowland, F. S., *Am. Sci.*, 1989, **77**, 36–45.
23. Dentener, F. J. and Crutzen, P. J., *J. Geophys. Res.*, 1993, **98**, 7149–7163.
24. Molina, M. J., Zhang, R., Wooldridge, P. J., McMahon, J. R., Kim, J. E., Chang, H. Y. and Beyer, K. D., *Science*, 1993, **261**, 1418–1423.
25. Rowland, F. S., Warg, C. and Blake, D., Proceedings of International Symposium on Global Cycles of Atmospheric Greenhouse Gases, Sendai, Japan, March 1994, pp. 38–51.

ACKNOWLEDGEMENTS. I thank Prof. B. N. Raghunandan and Prof. J. Chandrasekhar for encouragement and support. Dr C. Jackman, Dr A. Thompson, Dr P. K. Bharthia, Dr D. Wuebbles, Dr D. Fahey and Dr S. Madronich are thanked for useful correspondence and discussions. Dr L. A. Treinish is thanked for providing necessary tools for IBM data explorer to visualize the ozone data. The TOMS ozone data were provided courtesy of the National Space Science Data Center at NASA/Goddard Space Flight Center, Greenbelt, MD.

S. Parthiban, Department of Aerospace Engineering, Indian Institute of Science, Bangalore.

The 1995 Nobel Prize in Medicine: A tribute to the power of formal genetics

One of the most remarkable areas of achievements in biology during the past few decades has been the elucidation of the mechanisms that an apparently unstructured egg cell uses to transform itself into a complex-patterned multicellular organism. The first glimpse of this genetic blue-print was obtained through genetic studies with the fruit fly *Drosophila*. Fortunately, these results were obtained at a time when the recombinant DNA and molecular biological techniques were spreading like wild fire and therefore, even these esoteric genetic studies with *Drosophila* attracted the imaginations of biologists working with all kinds of organisms. Very soon the commonality of organization of all living organisms was reaffirmed and scientists could hope to decipher the genetic programme that controls development of as complex an organism as *Homo sapiens*.

Therefore, the committee for this year's coveted Nobel prize in Medicine has very rightly selected three *Drosophila* geneticists for their pioneering work on the genetic control of embryonic development and differentiation using simple but powerful tools of conventional genetic analysis. They are Edward B. Lewis of The California Institute of Technology, USA, Christianne Nüsslein-Volhard of the Max-Planck Institute for Entwicklungsbiologie, Tubingen, Germany and Eric Wieschaus of Princeton, USA. Lewis has been working with *Drosophila*, mostly by himself, for more than 50 years and published[1] a summary and analysis of the data collected by him over many years in his well-known *Nature* paper in 1978; Nüsslein-Volhard and Wieschaus published[2] their seminal paper, also in *Nature*, in 1980 when both were at the EMBO Laboratory in Heidelberg, Germany. These two papers have changed the course of contemporary biology in more than one way by allowing a new look at the transformation of a seemingly structureless egg into a complex, patterned and organized organism. The concepts generated in these papers have found very wide applications in studies dealing not only with animals but plants as well. It is remarkable indeed, that these two papers were based on simple methodologies of 'pure' or formal genetics with no 'sophisticated' or 'advanced' molecular biological techniques being employed. As is the wont of geneticists, they simply obtained a large number of mutations, selected those that affected early development, mapped the mutations on linkage maps and characterized the consequences of either individual mutations or specific combinations of the different mutations on the developmental phenotype of the individual. The only 'advanced' analytical techniques employed by these scientists were scanning electron and/or dark-field microscopy! It was the systematic analysis and a foresight in rationalization of the phenotypic effects of the various mutant genes that led them to formulate general principles of Nobel prize-winning consequence.

Study of embryonic development has fascinated biologists for a long time. Curiosity to know how a single-celled egg develops into a complex organism led to the growth of the whole field of Embryology to describe the morphological and anatomical changes taking place in a developing embryo. In biology it is often necessary to examine the abnormal so that the normal may be understood. Thus a conventional experimental embryological approach to the study of a complex process has been to disrupt one step in the chain of events and follow its consequence. This path was followed by generations of eminent embryologists during this century and their efforts did help in providing some

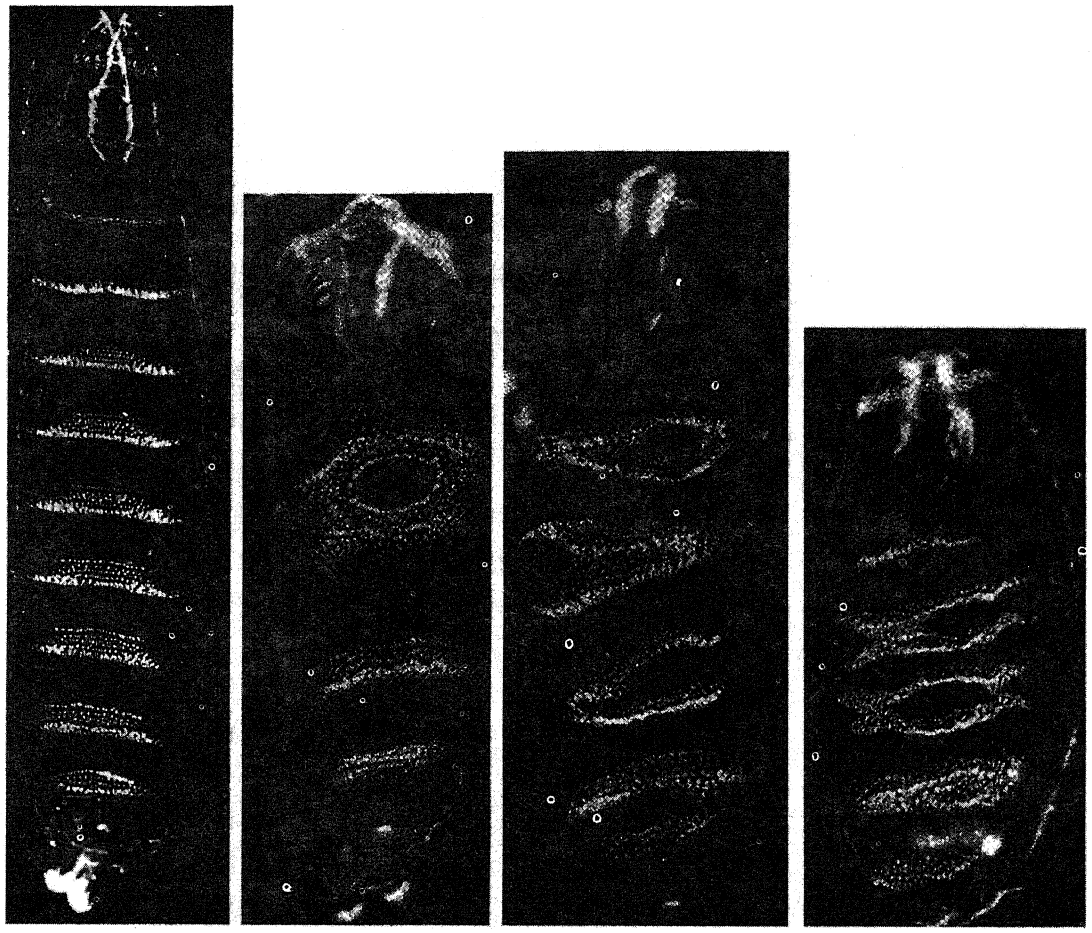


Figure 1. Examples of disruptions in normal segmentation of late *Drosophila melanogaster* embryos due to mutations in some of the segmentation genes identified by Nüsslein-Volhard and Wieschaus: the first panel shows the segmentation pattern in a normal late embryo; the second is an embryo mutant for the *krüppel* gap gene (with the prothoracic to 5th abdominal segments missing), the third one is mutant for the *paired* pair-rule gene (with the posterior compartment of all odd-numbered and the anterior compartment of all even numbered segments missing) while the last one has a mutation at the *engrailed* segment-polarity gene (with posterior compartments being affected). Identification of such defects in segment patterns in mutant embryos led the scientists to interpret the functions of the wild-type genes in normal pattern regulation.

cause–effects relationships in the complex events taking place during embryonic development. However, none of these approaches could address the central question of what molecules actually caused a particular cell in the embryo to decide or to know that it belongs to say head or tail region of the future animal or the leaf or root of the plant. With the growing realization that like all other things in living organisms, the development was also under genetic control, a geneticist would approach this issue by inducing random mutations in genes and identifying, not only the gene that has mutated but also the consequence of that mutation on the phenotype of the individual. It is this approach that led the three Nobel-prize winning scientists to provide the first insights into the complex signalling and regulatory events that 'tell' the undifferentiated cells to be committed to and follow a particular path of differentiation during early embryonic development.

To appreciate and understand the significance of these two papers[1,2], we need to briefly look at the way the *Drosophila* embryo develops (it is interesting to note that the basics of *Drosophila* embryology were properly understood only because of these and other subsequent papers, yet to understand these papers *today* we first need to look at what was historically learnt later: this highlights the farsighted intuition of these scientists as well as the power of genetic analysis). Unlike any other well-studied case of embryonic development, the *Drosophila* embryo follows a very unusual pattern of cleavage divisions during which only the nucleus divides repeatedly without any cell division; this results in a syncitial blastoderm stage which quickly cellularizes. This is followed by the demarcation, along the anteroposterior axis, of the repeating units of a defined number of body segments (parasegments) with the visible segment boundary bisecting each parasegment. As a result, each anatomical segment includes the posterior half and the anterior half of consecutive parasegments. Depending upon the antero-posterior and dorso-ventral coordinates, cells in each half (compartment) of the segment are committed to generate specific adult structures. A variety of genetic studies had established that even at the syncitial blastoderm stage, much before the morpho-

logical segments become apparent, each nucleus 'knows' its location in the embryo and also what its 'fate' is going to be when the cells begin to differentiate. The time gap between this 'commitment' and 'differentiation' can in fact be several days long. The basic question of how the nuclei in the syncitial blastoderm of *Drosophila* 'know' their location and their fate was addressed through genetic studies.

Although Lewis worked much earlier than Nüsslein-Volhard and Wieschaus, the genes that Lewis studied are called upon their duties later in embryonic development than the sets of genes described by the other two. Nüsslein-Volhard and Wieschaus argued that genes that control early embryonic development (and thus affect the critical determination events), if mutated, should cause malformed embryos and therefore be embryonic lethal. Therefore, they generated, through conventional mutagenesis procedures, a large number of random mutations in *Drosophila* genome and selected those that led to death of embryos. A simple microscopic examination of these dead embryos helped them determine the defect in the normal developmental pattern and thereby decipher the role of the mutated gene in the process (see Figure 1). In summary, Nüsslein-Volhard and Wieschaus showed that three sets of genes are sequentially activated after fertilization: i) the 'gap' genes define by their expression three major domains (anterior, middle and posterior) along the length of the embryo so that a mutation in any of these genes resulted in the loss of a contiguous series of structures along the antero-posterior axis; ii) the 'pair-rule' genes, working after the 'gap' genes, commit the nuclei to one of the 14 repeating segments along the length of the syncitial blastoderm: these were called 'pair-rule' genes since they seemed to work in alternate segments and a mutation in any of these genes caused alternate, odd or even-numbered, segments to be defective, and iii) the 'segment-polarity' genes which finally commit a blastoderm nucleus to the anterior or to the posterior compartment of a given body segment so that a mutation in any of these genes disrupts the polarity in each segment. These three sets of genes thus were shown to 'instruct' the blastoderm nuclei their polar co-ordinates and thereby define segmentation of the body, much before we can actually see it.

A parallel development around this time was the identification, again through mutational analysis, of a number of genes that were active during the period of oogenesis in the mother and whose products were later shown to be very orderly localized in the egg cell; the products of these maternally acting genes behaved like the classical morphogens to directly activate the different gap genes in different domains of the fertilized egg and through them the later acting pair-rule and segment polarity gene cascades.

The genetic system elucidated by Lewis comes into play after the above three sets of genes have carried out their work. Lewis worked with a very interesting family of mutations, designated by *Drosophila* geneticists as the 'Bithorax-complex' (BX-C): this name was given since the flies that carried such mutations typically showed duplication of the thoracic region (see Figure 2). Such mutations belong to an interesting class, the homeotic mutations, a term coined by W. Bateson[3] in 1894 for malformations that substituted the pattern of one body region for another. Very early in the growth of *Drosophila* genetics, a number of homeotic mutations (e. g. the *bithorax, aristapedia, proboscipedia*) were identified. The existence of homeotic mutations signified that certain genes 'tell' the cells at some stage of development as to what they are destined to become and if, as a result of mutation, these genes gave a wrong instruction to a cell, the cell develops a different structure at a wrong place (the homeotic transformation). Thus such mutations provide an unique opportunity to ask what these genes 'tell' the cell and at what time. Lewis approached this goal by analysing the large number of mutations that mapped to this locus and which transformed a thoracic or an abdominal segment of the fly into another in a complex but predictable fashion. He had earlier shown that this locus is a complex locus being composed of several discrete genetic units (hence these individual units were initially termed 'pseudoalleles') which were very closely linked but each had a distinctive effect on the phenotype of thoracic and the abdominal segments. He summarized an enormous wealth of his own experimental data in the 1978

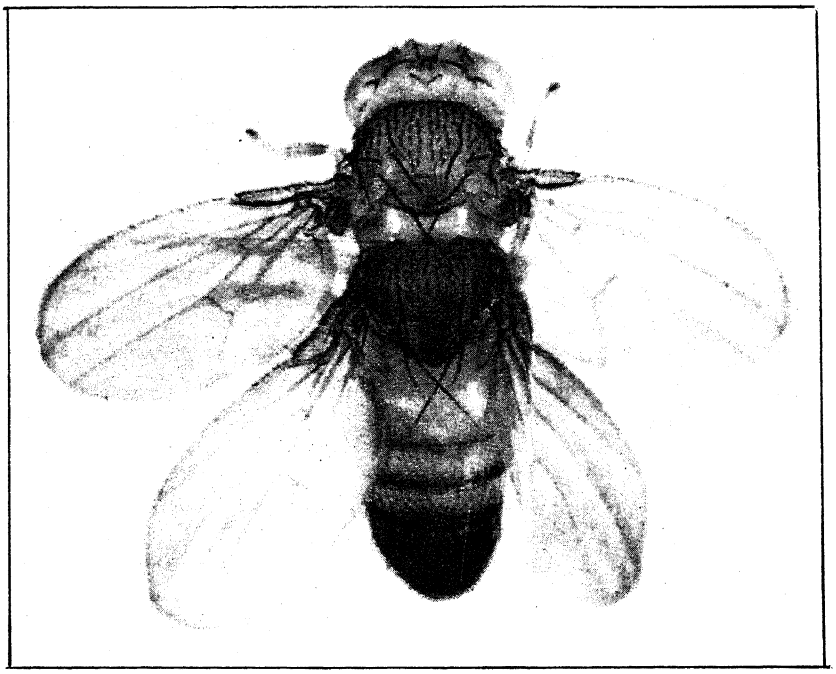


Figure 2. A typical four-winged fly generated in Lewis experiments when certain *bithorax* mutant alleles were brought together in heterozygous combinations: such a combination of the mutant alleles lead to a homeotic transformation of the meta-thorax into meso-thorax so that the normally reduced second pair of wings (the halters) in the meta-thoracic segment now develop into full wings like those in the meso-thoracic segment.

review published in *Nature*[1] to show that the different genetic elements comprising the BX-C control most of the body plan of the fly. He showed that a chromosome that had lost almost whole of the BX-C resulted in an embryo with all its body segments being similar (no abdominal segments are formed) and that addition of the different parts of the BX-C one by one can restore each segment in a defined order. Lewis argued that his results demand that the BX-C contains a minimum of eight genes that 'code for substances controlling levels of thoracic and abdominal development'. A most remarkable point noted by Lewis was that the order of these genes in the BX-C was collinear with the structures along the antero-posterior axis of the fly controlled by them. Lewis also argued that each subsequent posterior segment of the body recruits the services of an additional genetic element in the BX-C linear array. Subsequent studies (see ref. 4) at molecular and genetic levels, however, revealed that Lewis misinterpreted the *cis*-acting regulatory elements in the BX-C for structural protein-coding genes: the BX-C actually has only three structural genes but multiple *cis*-acting regulatory elements which are recruited in a serial fashion to activate one or more of the structural genes of the BX-C, each of which can produce more than one kind of message due to alternative processings of the primary transcripts. However this misinterpretation does not in any way undermine the far-reaching implications of his conclusions that the BX-C genes specified the developmental pathways to build the characteristic pattern of the different body parts and that these genes worked locally. The BX-C and other homeotic genes (like the Antennapaedia-complex or ANTP-C) are also called as the selector genes which take over from the segment-polarity genes and select the specific structures that each compartment has to elaborate as the embryo continues its development[4].

While working with BX-C, Lewis did face intellectual opposition from contemporaries since the genetic organization of BX-C defied the then held beliefs about a 'gene': some mutations mapping to this locus suggested that the whole system of the *bithorax* genes was integrated while others pointed to existence of separate and separable elements of function. Unmindful of criticisms and scorn of some contemporary molecular biologists, Lewis pursued the task of understanding the system of bithorax genes which resulted in his great 1978 paper. An interesting point mentioned by Peter Lawrence in his introduction to the historic 1978 *Nature* paper by Lewis in a collection of 'Outstanding Papers in Biology'[5], regarding the publication of this paper is worth noting. To quote Lawrence[5], 'A publication from Ed Lewis is an unusual and special event. In the past 40 years his papers have been few, and their originality and idiosyncrasy has found them natural homes in unrefereed conference reports. In 1978, I had a "phone call from him to ask for help in submitting a new paper to *Nature* – an article which would consist of a 'review' of a vast amount of Lewis' own unpublished work". Peter Lawrence urged the *Nature* to 'snaffle' the paper without question and to publish it unrevised, lest the quibbles over details may cause the paper to be sent elsewhere or even delay it for a decade! As Lawrence further notes 'it is just the kind of contribution that the modern system of reviewing, with its rigid conventions and predilections against imaginative speculation, would censor. We should all be glad that *Nature* decided to publish it.'

How did these papers dealing with 'pure' genetics stimulate such a global interest? In 1978, when Lewis had published his paper, few could appreciate the import of his speculative conclusions (e. g. William McGinnis, one of the discoverers of the 'homeobox' (see later), notes[6] that the Lewis paper 'is oft cited but rarely read in its complex entirety'). Fortunately for Nüsslein-Volhard and Wieschaus, the scientific community in the early eighties had become receptive to prospects of genes controlling development and also this community was now ready to look at the genes from the molecular point of view. Combining the powers of 'formal' genetics with recombinant DNA technology, it was soon possible to clone some of these 'developmental' or 'segmentation' genes and to ask more directly how did they function. Consequently, it was possible to understand the biochemical bases for the observed effects of the various segmentation gene mutations characterized by Nüsslein-Volhard and Wieschaus and to verify the inferences drawn by them from their mutational analysis.

The most impact-making discovery that followed from these papers in the early eighties was the finding of a short stretch of highly conserved DNA sequence in the coding regions of a number of homeotic and segmentation genes. This sequence, aptly termed the 'homeobox' (denoting its association with the homeotic genes), provided a plausible biochemical function for the action of the homeotic genes whose products now appeared to be DNA-binding transcriptional regulators. Discovery of the 'homeobox' was a prelude to a flood concerning homeobox genes and homeodomain proteins, a flood that has channelled into a steady river of homeo-publications, fed by many tributaries[6]. Quickly, the homeobox was making its appearance in genes from as diverse organisms as man, mouse, chick, frog, sea urchin, nematodes and even plants (see ref. 6). Its high degree of conservation provided a simple handle to clone the developmental genes from a large variety of organisms and for many a biologist (as well as the common man since this topic had become the subject of newspaper and magazines in the eighties), the homeobox was (somewhat mistakenly!) equivalent of the Rosetta stone, the universal genetic key to body plan. The most remarkable fact that emerged from these studies was that the organization of the homeotic genes like the BX-C and ANTP-C in *Drosophila* and their counterparts in other organisms (including man) was identical.

These studies thus led to the emergence of the fundamental concept that in all organisms, the events of differentiation and pattern formation result from cascades of activity of specific groups of homologous genes which sequentially 'tell' the cells of their spatial and temporal co-ordinates and accordingly each cell gets committed to a particular path of development. Most of these genes produce transcription factors that regulate the expression of other genes. These studies showed that, as in any language a limited set of alphabets are used to create an endless variety of words and expressions by varying their combinations, in the biological system also, a limited number of regulatory genes can be used to generate a large array of patterns by expressing in different combinations in dif-

ferent cells along the antero-posterior and dorso-ventral axes.

Ever since Thomas Hunt Morgan introduced the fruit fly or *Drosophila* to genetics early in this century, this small insect has been central to the growth of the subject of genetics and thereby to many fundamental concepts in biology. It is interesting to note that Morgan himself was an embryologist who turned to genetics with a view to understanding the central issues of development and differentiation. Although by 1913 Morgan gave up 'efforts to deal simultaneously with genetics as transmission and genetics as development'[7], it is interesting to note that Morgan's objective of understanding development in terms of genetics has finally been achieved and the power of this approach been duly recognized by this year's Nobel prize in Medicine.

1. Lewis, E. B., *Nature*, 1978, **276**, 565–570.
2. Nüsslein-Volhard, C. and Wieschaus, E., *Nature*, 1980, **287**, 795–801.
3. Bateson, W., *Materials of the Study of Variations Treated with a Special Regard to Discontinuity in the Origin of Species*, Macmillan, London, 1894, p. 85.
4. Lawrence, P. A., *The Making of a Fly*, Blackwell, London, 1992.
5. Lawrence, P. A., in *Outstanding Papers in Biology*, Current Biol. Ltd., London, 1993, p. 93.
6. McGinnis, W. A., *Genetics*, 1994, **137**, 607–611.
7. Falk, R. and Schwartz, S., *Genetics*, 1993, **134**, 671–674.

S. C. Lakhotia, Department of Zoology, Banaras Hindu University, Varanasi.

CORRESPONDENCE

If education is important, so must be the teachers!

I applaud the Indian Academy of Sciences' effort of exercising themselves on an important aspect of education in our country. In fact it is basically the stony silence of the 'intelligentsia' of our country that has left most of the matters that concern our lives in the hands of politicians, bureaucrats and judiciary. The result is there for all of us to see. The Academy's effort is thus very welcome.

I believe, however, there is need to look at another essential, complementary question, *i.e.* what sort of facilities should be provided to the teaching profession so that teachers become effective instruments in translating good ideas into reality.

Two major causes of the progressive deteoriation of education in universities in our country are:

(i) teaching has never been given a place of pride in hierarchy of jobs in terms of salary and service benefits,

(ii) there is rampant corruption in educational institutions in matters of selection and promotion of teachers, organization of research, etc.

The following suggestions are worth considering:

(1) An Indian Educational Service (IES) for college and university teachers and equivalent state-level educational cadres for school teachers be created, making it the BEST public sector job in terms of salary and service benefits. The necessity of making such jobs the best in the country can hardly be overemphasized if the nation really feels that imparting education and science education in particular is of utmost importance for its growth—more important than handling bureaucratic and law-and-order problems.

2. The pay-scales of teachers at all levels be time scales whereas promotion to higher cadres be made on periodic evaluation of their work by the peers drawn from as distant quarters as possible. Simultaneously, new posts be created for fusing young talented people from outside. This is, in fact, the procedure followed in research institutes and to some extent in colleges but not in the Universities! The present practice of putting people into available posts in the Universities has led to immense intra- and inter-departmental anomalies and only encourages servitude, favouritism and corruption at the cost of sound work-culture.

A minimum set of criteria, generally agreed upon, be insisted as requisite for promotion to various cadres. This should also include the students' evaluation of the teachers to be sent directly and confidentially to the evaluating authorities, say Vice-Chancellors in case of the Universities and inspectors in case of schools.

The whole point is that a teacher should be assured of his growth in the cadre only through his positive contributions to the system like teaching, research, commitment towards social issues, taking care of the overall growth of student life in the institutions, etc.

We should all accept that despite the likes of Raman, Ramanujam, Bose, Saha and Bhabha or various science institutions of repute, science has not really become a part of our culture. Our culture still continues to be dominantly feudal and colonial, denying all that science in spirit and form stands for. Resurrection is only possible through a sound educational system and that in turn requires a competent class of teachers at all levels taking pride in their profession. This requires a system that we lack so far. This is where Governments and intellectuals can come in to help erect it. Good players in good playing conditions will, by definition, make a good game. Can we ensure it ?

LAMBODAR P. SINGH

Department of Physics,
Utkal University,
Bhubaneswar 751 004, India

COMMENTARY

India's scientific community: A cultural crisis?

A. N. Mitra

Over the last 15 years, the state of science in the country has been under increasing scrutiny at major academic fora – the Indian National Science Academy[1], Indian Academy of Sciences[2], the Indian Science Congress sessions[3], periodic DST status reports[4] and Seminars in the Society for Scientific Values[5]. The theme has almost invariably been the same: There is little qualitative impact of Indian Science on the international field despite proliferation in the number of papers published in various journals. This, in turn, is a direct reflection on the state of 'University education in science' in the country. The preamble for such discussions has always been predictable: What has gone wrong since Independence in a country which can rightly boast of the pioneering achievements of Bose, Raman and Saha, all done indigenously in this very atmosphere? And let there be no mistaking the extreme poverty and the consequent non-existence of the so-called 'facilities for scientific research' prevailing in those times, and which the modern generation of academics and professional workers never tires of complaining about. Indeed, the great achievements of these giants in the pre-independence days, sitting in this very soil, as distinct from the equally impressive successes of their 'provasi' counterparts (Bhabha, Chandrasekhar), would even tempt one to wonder if there is any other 'mystery' apart from 'funds' which lies at the root of 'success' in a scientific endeavour. Ramanujan was of course a 'class' by himself, and cannot be understood in 'rational' terms. One must, of course, hasten to make a prior allowance for a fortuitous circumstance which existed 70 years ago, *viz.*, the state of flux that science, especially physics, was undergoing in those days when it was perhaps easier for a single individual to make a decisive impact on the international map than it is today. Yet a closer scrutiny of the 'nature' and 'thrust' of their investigations strongly suggests the influence of an 'indigenous' culture that is quite independent of its western counterpart, despite the general background of Western Science, in which these investigations had been carried out. This is especially true of Bose's great discovery which one may venture to guess would not have been easy to make in the then prevailing atmosphere of Western Science which had not (and still has not) recovered from the strongly 'deterministic philosophy' of the nineteenth century, and was very much 'ill at ease' with the new paradigm that had been thrust on it by half a dozen western minds in quick succession (Planck, Einstein[2], Rutherford, Bohr and de Broglie). (In this connection I am reminded of a remark by Abdus Salam in 1970 (ref. 6), in the context of a discussion on 'symmetries', namely, it is often easier and more natural for an easterner to appreciate the significance of concepts based on apparently 'conflicting' ideas such as particle–wave duality than his western counterpart, and indeed went on to cite Bose's great discovery of the very concept of indistinguishability as an example).

Apparently there are vital ingredients (other than funds and facilities) in the indigenous cultures and traditions of a given community that go to make up its scientific ethos and values.

Pre-independence scenario

Now to come back to the question which has been bothering our academics[1–5], it is necessary that one must not fight shy of looking back to the pre-independence scenario for comparison. As long as the country was under the foreign yoke, the political atmosphere was such as to keep the academics and intellectuals confined to the portals of the modest university system for all it was worth. The foreign government was simply not interested in making use of our scientific talent for genuine developmental purposes (except perhaps for occasionally hiring some scientists for help in their war efforts). Nor were the members of the scientific community anxious on their own in offering their services to the then government against the emotional tide of the freedom movement. The safest abodes for their intellectual pursuits under the circumstances were thus necessarily the educational institutions which also 'housed' the sciences in a modest way in the absence of alternative, exclusive, abodes of scientific pursuits (with rare exceptions like the Indian Association for the Cultivation of Science, the Bose Institute in Calcutta and the Indian Institute of Science in Bangalore). The result: whatever little (?) output came out of this system was truly indigenous and this included the signal contributions of the giants noted above.

The 'culture' of these educational institutions in the pre-independence days was also quite different from what it is today. The 'ambitions' of the teachers were confined to their modest academic pursuits which included offering their best attention to their students, as if they were their own children, and not from any sinecure motives. (I have seen my late father (Jatindranath Mitra) spending hours together with his students at home, over and above the limited hours spent with them at his Ramjas College, with the sole incentive if any being one of 'competition' with sister institutions for his students' performance at the University examinations, as if their successes were his own! Any other motive, especially the 'tuition' motive which is so prevalent today, was simply unthinkable then). No wonder, the students got their academic dues without ever asking for them, but more importantly, this kind of gurukul atmosphere had a profoundly salutary effect on their moral fabric which they were to propagate later to their progeny. The seminal researches of Bose–Raman–Saha were already ample evidence of the modest but indigenous research atmosphere present in the then University educational system pioneered by great visonaries like Asutosh Mukherjee (Calcutta), Madan Mohan Malviya (Banaras) and the Maharaja of Mysore, all of whom took great personal interest in bringing together

bright young workers and encouraging them in their efforts.

Post-independence: Decline of university system

This 'garden atmosphere' was however too Utopian to survive the economic challenges of the post-war era which also happened to coincide with the post-Independence era. With the dawn of Independence, our founding fathers rightly stressed the need for rapid industrial development, seeking to compress as it were the two centuries of Industrial Revolution in the West to at most a few decades. This new emphasis called for fresh priorities with its inevitable impact on the society. In this review, we are however concerned with only some aspects of this impact on a small but vital section of the society which is supposed to nurture science through a sound educational system. In the new dispensation of the then government, the universities found themselves suddenly reduced to the status of 'second class' institutions, as if to contradict the very noble sentiments of no less an authority than Jawaharlal Nehru on the true role of a university in a free society as the symbol 'for compassion, for reason, for the adventure of ideas and generally for the onward march of the nation towards still higher objectives[7]'. Not that there was any deliberate attempt to downgrade the universities, but in the then nascent atmosphere, their social status registered a sharp decline in the face of a new 'cultural revolution' caused by the creation of a chain of government-sponsored organizations such as the CSIR, ICAR, ICMR, AEC and the various national laboratories under their respective jurisdictions on the pattern of their British counterparts. The external symptom of this development was a sudden imbalance caused by the outflow of talent towards these new pastures at the immediate cost to the educational institutions which had hitherto been their traditional abodes, as well as fountainheads for the supply of intellectual and scientific talent in any country, but the negative effect penetrated much deeper (see below).

Now the growth of government-sponsored scientific organizations is inevitable in any developing society and their creation on the British pattern cannot be faulted on intrinsic grounds. Indeed in the post-war era, a similar development was also noticeable in the French System under the CNRS and like organizations[8]. However, unlike in Britain or France where the firm tradition of liberal intellectual values has always ensured a steady supply of a sizeable fraction of the intellectual cream of the society within their university systems without fear of any major depletion of talent, a phenomenon of this nature proved much too harsh for the delicate educational structure of this country in the nascent atmosphere prevailing in the post-independence period. And to make things worse, there was no evidence of a corresponding degree of concern in this country for countering the impact of this new culture through suitable checks and balances. Instead there was just a one-way development: The discovery of new form of 'glamour' in these prestigious institutions in the Indian context tended to attract the scientific manpower to the new environment by way of better economic and social prospects, while on the psychological plane, it had a negative effect which should have been anticipated: A 'devaluation' of the University system through the implied perception of a sort of 'poor cousin' status for the latter in the eyes of the society at large. Unfortunately the government made no effort to introduce any corrective measures against such an obvious imbalance which was to have a long-term detrimental effect on their healthy growth in the new scenario. For on the one hand the University system now found a major rival claimant for the modest state support that it had hitherto been used to, with the greater part of the available funds for development being diverted to National Institutions now manned by people identified as the 'top scientists' in the country many by whom had little actual contact with teaching at the brass tack levels. On the other hand, the visible indifference of these new 'managers of science' towards the problems of their 'poor cousins' (presumably for lack of any direct experience with the working of these institutions) generated a psychological feeling of social inferiority within the university community, to such an extent as to generate a genuine desire among its smarter members to migrate from the university system to 'new pastures', e.g., 'abroad'. Thus, far from strengthening the fabric of the educational structure, the new policy of the government seemed to be tantamount to 'allowing' the traditional educational institutions to 'decay with time, when its long-term concern should have quite the opposite, *viz.*, a continuous generation of scientific talent which only the latter could provide for a steady flow of scientific manpower for the expanding organizational activities in science and technology. Incidentally, this 'decay' has been strictly at the qualitative level, for with the passage of time there has been a steady proliferation of the number of universities in the country, almost in tandem with the increase of U. N. membership during the same period (while the FIVE permanent members of the Security Council dictate the deliberations and decisions). This fact, together with the utter frustration of the teaching community due to the lack of basic amenities for work, as well as of incentives for promotional avenues over the last three decades, has probably been the biggest single factor responsible for the parlous state that the university system finds itself today. The peculiar form of 'trade-unionism' now witnessed in the educational institutions is hardly distinguishable from what one is normally accustomed to in the industrial sector, and is a far cry from their 'cloistered atmosphere' of half a century earlier. In the public perception, the university system has now acquired the dubious distinction of a 'second-class citizen', what with the so-called 'promotional farce' (especially in the professorship category) which has merely served to erode its credentials further. In all fairness though, the question of 'which came first, the hen or the egg?' remains at best a speculation.

Degeneration of scientific manpower

The above is only one side of the story, one accruing to a vital sector of the society, the university system, to the decline of the scientific enterprise in the country. The other side of the coin, the state of the scientific community which forms the bulk of the scientific manpower of the government-run institu-

tion, is no more rosy, if one has to go by the frequent concerns expressed by various Academic bodies[1–5], case studies by professional experts[9], as well as numerous expressions of public concern reported in various magazines and newspapers in recent times under titles like 'tragic gap' [ed. Indian Express 3.2.95], 'plight of scientists' [Times of India, 11.3.95], to mention a few samples. (This is hardly surprising, since after all, the 'timber' of scientific manpower in the country comes from the same stock as can be supplied by the educational system itself, irrespective of whether the initial advantage of the former was at the cost of the latter). The laments expressed in these media columns, including some[10] from members of the general public, are mere variants of a common theme: 'a moribund state of science in the country'. The symptoms span a whole range of factors, both specific and non-specific. Some items are[10]:

(1) A lack of sensitivity (of authorities?) in identifying and encouraging scientists;

(2) Failure of senior scientists to champion the cause of their junior colleagues;

(3) Public unawareness of the research-funding policies. As a result, huge grant sanctions are often based on political clout of the scientist in an atmosphere of corruption and sychophancy at the bureaucratic level;

(4) Lack of peer support of upcoming young scientists due to the reluctance of the top rung of scientists to encourage talented and ambitious youngsters (barring some genuine cases). This stems from a sense of insecurity (born out of their own inadequacies, lack of commitment and complacency). The resulting effect is a gradual extinction of the 'flames of scientific temper'[10] in the young minds through a denial of access to facilities, and/or only short-term support to their research.

(5) Callous attitude of officials in the Alma Maters of young scientists, resulting in delay of sanctions, bills settlements, etc, so that the finance and purchase sections decide the time table.

The common refrain of most of these complaints is a pitiful lack of 'funds and facilities', together with a very unjust form of distribution which leaves this meagre resource in the hands of a small but influential group which would resist any sharing of scientific power down the line. Underlying this refrain is an implicit belief that the amount of any worthwhile research output is directly proportional to the available *funds and facilities*, as if to suggest an exclusion of all other factors! This is certainly an oversimplification but it goes to show how deeply the malaise has penetrated the psyche of the average scientific worker. The kind of research that is directly proportional to the 'fund available' is still quite substantial in this country, as a recent discussion in one of the premier Academies[1] would suggest, but how much qualitative impact it makes on the international scene, is anybody's guess, despite our claims to the possession of the fourth largest scientific manpower in the world. Apparently some other vital element/elements conducive to the growth of a healthy science are also involved.

Role of 'national ethos' in the growth of science

What is this vital element/elements? Perhaps a single phrase in which it/they can be summarized is a sense of 'national ethos' which binds a great nation together in terms of its different components, scientists included. Indeed it does not take too much scrutiny of the scientific achievements from any great country which has made a sizeable impact on international science (UK, USA, USSR, France, Germany, Italy and Japan) to convince oneself of the 'sense of belonging' of its scientists as its nationals trying to project the work of their colleagues as if it were their own. And this stance is visible quite naturally and spontaneously, without the slightest effort at any 'pretence'. Any member of our scientific community who has interacted with his/her 'foreign' counterpart will not fail to testify to this common characteristic which binds them together, transcending the international character of science. One would also discern a broad degree of 'cultural similarity' as well as a sense of national solidarity in their respective publications (as revealed by their referencing patterns, among other things). This is not to say that scientific pursuits in these countries are free from problems, but they are of other types (also present in our system) and their identifications, and solutions are at least not hampered by the 'cultural identity' problem.

Let us now compare the corresponding scenario in our country. To keep the record straight, there is no question of a sense of national identity not existing at the socio-political level (in the face of any external threat, we are all one), but are we equally sure that the same degree of 'cultural solidarity' obtains within the subsystem of the scientific community (which after all is not threatened by an external enemy)? Despite possible protestations to the contrary, the answer seems to be an emphatic 'no'. And the evidence lies not just in the relatively innocuous fact of our scientists being more comfortable with their activities when they 'go abroad' with obvious attractions on both the economic and the 'fund–cum–facility' fronts (this is merely an external symptom, not a cause), but rather in their very attitude of indifference to a sense of cultural and ethnic togetherness compared to their more successful western counterparts. This attitude is more easily noticeable in the themes and referencing patterns of their publications which evince an unmistakable tendency to 'lean' towards western thoughts even for mundane and routine ideas, and often show a tendency to ignore similar work emanating from their compatriots, in favour of corresponding references from more 'visible' western colleagues. This may well be due to a feeling of insecurity born out of a lack of confidence in indigenous research output, but it has the effect of snowballing into a peculiar 'collective culture' in which the indigenous component becomes the victim. It may be argued that since basic science is international in character and there is only one accepted theory by and large, such an attitude is unexceptionable *vis a vis* the theme of a paper. While this is a priori a very powerful argument, it is not free from limitations, for the same logic should also apply to the various other nationalities in western science. Yet one invariably notices some distinct 'nationalistic' flavours (American, Russian, Italian, Japanese, etc.) in their respective publications, despite a common pursuit of international science. The only sensible answer to this paradox lies in the simple fact that within the broad framework of 'accepted sci-

ence', there is plenty of scope for diversification of ideas designed to address limited sectors of investigation, and it is through such possibilities that one often arrives at 'new' insights on an 'old' problem. There is good reason to believe that such 'insights' have a sort of 'regional character' reflecting the 'national ethos' of the investigators concerned. And in any case it is difficult otherwise to account for the fact that most original discoveries have emanated from one or more of these *individual* countries, with their clear imprints, and very rarely on a 'multinational' basis.

To illustrate the nature of these 'insights', one invariably finds an abstract (often mathematical) flavour in Russian and French research publications, in contrast to a more practical 'applicational' orientation in their British and American counterparts. German and Italian Physics which used to be oriented towards the abstract and applicational respectively in the pre-war days, have undergone drastic changes since then, with a certain degree of intermixing of the two, probably as a result of the traumatic effects of the War on their respective social structures. The Japanese ethos, on the other hand, shows an interesting transition from the Eastern abstractness of the pre-war physics (Yukawa, Tomonaga and especially the Nagoya Group) to an applicational (bordering on engineering) orientation, especially since the War, perhaps symbolizing a *collective national effort* for 'catching up' with the western nations on the technological front; yet the key element has been unmistakably a collective nationalistic ethos.

Against this backdrop it is interesting to ask if any collective 'nationalistic ethos' is noticeable in the Indian context, even after half a century of the post-independence period. The characteristic imprint of 'Eastern Thought' was clearly visible in the pioneering researches of Bose, Raman and Saha, carried out in the 'splendid isolation' of their respective environments. It would be idle to ask if there was evidence of any 'national identity' under the shadow of a foreign ruler, but the themes of these papers were 'characteristically Indian' which did not prevent them from having a decisive bearing on the international frontiers of physics. True, it was a period of upheaval in Physics when it was getting reshaped in tune with the new realities, and when the chances of making a significant impact due to individual efforts were a maximum. However, it is interesting to ask if a 'Bose' or a 'Raman' could have made a similar discovery in the present environment of science in the country after 50 years of independence when our 'Eastern' philosophy should in all fairness have crystallized into a well-defined 'national ethos in science', as in the more 'mature' countries of the west. One major problem he would face is the enormous growth of Physics since the consolidation of Quantum field theory (though the current state of this theory is again in a state of flux !). The more serious handicap for him would be cultural: he would find himself as a 'Rip Van Winkle' with no trace of the indigenous scientific culture that he had (perhaps inadvertently) *initiated* 70 years ago, in today's 'voo-doo' atmosphere which pervades every section of the society, science included, making it virtually impossible for any well-defined indigenous culture, scientific or otherwise, to evolve. True, in the new world order some degree of intermixing of cultures of the East and the West is inevitable, yet it is fair to say that by and large, the cultural identities of the major nations have not been lost in the process.

Decline of cultural values

So, where do we stand in this new scenario? More pertinently, how did this remarkable increase of entropy in our 'science culture' come about? While part of the reason must be attributed to the worldwide phenomenon of increased mobility in the post-war period which has led to a considerable intermixing of scientific and cultural values, one suspects that there must have been other important factors without which it would have been difficult to account for such a qualitative transformation in a country of India's size and diversity which has maintained its cultural identity over centuries. One major factor that may stand some repetition in retrospect is the government's (short-sighted?) perspective for 'development in a hurry' by indiscriminately drawing on the Nation's 'intellectual reservoir' (the Universities) without a balancing concern for the long-term effect on the cultural health of the nation. This was done by declaring its dependence on the scientific community to be sure, but the actual implementation came about through a hurried 'patronage' to a few 'top scientists' (chosen arbitrarily) in a 'nascent' atmosphere where the Society's perception of a healthy balance between its scientific manpower and a sound educational base to generate it had not had time to crystallize. Now, while the choice of a 'Bhabha' to initiate the Nation's atomic energy programme would be deemed unexceptionable even after 50 years today, the arbitrary choice of many others as 'science managers' led to a visible imbalance in the educational and scientific community then at a crucial stage of evolution. Among various fallouts, a 'competitive' phenomenon was discernible, with some senior academicians leaving their natural habitats (*viz.*, the universities) for similar (lucrative?) placements with the government. From the long-term perspective of a healthy growth of the educational system in the country which is the fountainhead for the supply of sound scientific manpower on a continuing basis, such an episode, right at the dawn of independence, had a profoundly negative effect on the traditional image of an 'ideal society' normally associated with the pivotal role of the 'universities' in projecting this noble image. Indeed, in the eyes of the academic community, especially of the younger generation, this transformation which put the glamorous image of the Science Manager above the low-profile picture of the 'teacher' working in an educational institution was tantamount to a repudiation of its traditional value system, the long-term effect of which may be partly captured by the spirit of the following poem due to Robert Browning (?) on the occasion of William Wordsworth's appointment as Poet Laureate of England:

'Just for a handful of silver he left us;
Just for a riband to stick in his coat.
Got only one of which fortune bereft us;
Lost all others she lets us devote.
They who had gold to give doled him out silver;
How all our copper had gone for his service.'

In the Society's perception, such a lopsided perspective at the very dawn of

independence, without a corresponding corrective measure to ensure that no damage was caused to the long term health of the educational system, set the tone for *defining* the relative priorities for development of the universities *vis a vis* governmental organizations, and indeed repudiated in the process the lofty ideals set out by Jawaharlal Nehru himself[7] on the role of the universities in building up the mental health of the Nation. Once the relative priorities were so defined, the subsequent evolution of the university system (*vis a vis* the scientific organizations under government charge) took its natural course where 'shortage of funds' for scientific education was but a symptom of the many consequences to follow. And if the then Society accepted this 'value system' without proper scrutiny of its long-term effect on itself, (since in the ultimate analysis the health of the scientific manpower would be determined by that of its basic source–the university system), it is merely the price of a costly mistake made half a century earlier that we are paying today in more ways than one.

Perhaps the biggest casualty of this lopsided value system was its failure to develop some well-defined 'schools of thought' with a characteristic Indian flavour, for which the best breeding ground is the university system itself. Of course, there had been attempts to develop such schools on a localized scale, but these pockets did not live long enough to grow into an all-embracing 'Indian ethos' of the type discernible in the more successful western societies noted above. With the 'individualistic ethos' that have characterized our scientific community (in company with the rest of our countrymen, as Gandhiji had recognized well enough to advocate basically small scale industries!), the more 'influential' among them merely took advantage of this lop-sided value system to promote their individual career interests through lobbying with the appropriate authorities in a well-defined hierarchy (Science managers for the middle level scientists; the political authorities for the top rung). Apparently, the sole criterion of scientific 'visibility' lay in some sort of State recognition (in the form of national awards, higher placements, etc). Others with more conservative forms of ambition for visibility in science often chose to look for more suitable opportunities abroad. And all this transition occurred before the scientific community in the country had reached a critical size, not just in numbers but in quality, which is a standard bulwark against any major scientific injustice or aberration in any 'mature' nation. No wonder, the degeneration into such an 'opportunistic atmosphere' was not conductive to the evolution of a holistic scientific culture, and the same got progressively reflected in more ways than one. A major signature has been the publication pattern of our scientists which have shown a strong tendency to routinely follow western ideas howsoever trivial, even in mundane theoretical investigations, while in the experimental areas of applied science the themes have often no relevance to the needs of their local environments. In either field there is little sign of effort to develop an indigenous culture as is normally understood in a western science community. Instead one witnesses a hybrid 'value system' in which a desperate *cultural* dependence on 'Western' science for research ideas has come to be matched by an equally compulsive culture of sychophancy with which the members of the upper echelons of the scientific community 'look' to the 'authorities' for security of their own positions of power and influence, while those in the middle levels in turn, are only too eager to pay court to their 'superiors' for sheer survival. This typical scenario, which is reminiscent of the political culture in the country as a whole, has got perpetuated in the failure of a strong 'scientific constituency' to develop as a natural custodian of moral and scientific values in any country. Such a body which evolves gradually in a mature educational and cultural base, did not have enough time to crystallize in a healthy manner in the atmosphere prevailing in this country. We now have a 'Society for scientific values' which periodically 'assesses' the moral and cultural health of the scientific community, but it comes too late in the day (and with too little mandate to do anything but scratch the external surface of the malady) for a meaningful impact on the health of this community to day.

Strangely enough, the hazards of too much dependence on the western environment have lately started telling on the more ambitious among Indian scientists who have been experimenting with a scientific career abroad. An interesting study by an eminent scientist and his colleague living abroad[11] reveals that after the first two stages (pre- and post-doctoral) in US universities and scientific institutions which the young Indian worker finds extermely satisfying for his knowledge, and rewarding for his career, there begins a long period of professional frustration for the more ambitious among them who have an urge to compete for professional recognition with their western counterparts. No wonder many of these NRIs, including scientists, are now anxious to come back to play an active role in the Indian scenario (including the setting up of an 'ideal' science university ?), perhaps after tasting the limits of 'intellectual satisfaction' available from the West, *vis a vis* their personal ambitions. It is precisely in such situations that the true significance of a characteristic (national) ethos, something that is naturally present in the western system but has somehow got short-circuited within the Indian scenario in the process of 'evolution in a hurry', comes out in bold relief.

To keep the perspective right, it will be a gross oversimplification to claim that Indian Science has scored no visibility at all since independence, for the Nobel recognition of Subramanian Chandrasekhar a few years ago stands out loud and clear, despite the technicality of the citizenship anomaly. Among other pioneering achievements in recent times which have helped put India on the world map are those of G. N. Ramachandran (Triple helix), S. Chandrasekhar (Liquid crystals), A. S. Paintal (Science of breathlessness) and C. N. R. Rao (Surface chemistry) to name a few. Again, in the area of the theoretical sciences (mainly physical) there are a few institutions of science, mostly outside the university system, where good quality work has attracted 'proper recognition' in the international scientific community, in the sense of a "Conformity with 'normal standards' prevailing today in the international field", but nothing more concrete can be claimed in this regard. These are indices of better 'funding and facilities' in improving the quality of work in a standard institution whose

'cultural base' continues to 'look west' for fresh ideas. However, considering the vast potential of the scientific manpower in the country, a growth of much bigger dimensions would have been expected. This brings back the bigger question of the possibility of cultivating a more indigenous scientific culture (for which the intellectual 'software' can come only from our own educational system) and which alone can hold out the prospects of generating 'fresh ideas' indigenously without the compulsion of 'looking west'.

Indigenous scientific 'hardware': AEC and ISRO

For a comparison of this scenario on 'cultural software', with the corresponding scenario on the technological front, it is useful to keep on record that at least one area in which a centralized command (under an ideal leader – Homi Bhabha) has paid off quite handsomely is a successful development of an indigenous culture in the atomic energy research programme which has grown from strength to strength over the decades since independence and has now developed a strong infrastructure which is not only largely self-contained but has been continuously expanding its activities in several dimensions to generate appropriate spin-offs without external support. The international stature of this organization today owes much to the single-minded devotion of Bhabha's able followers like Raja Ramanna and his colleagues who have put this country on a firm footing on the world map. In parallel with the AEC, the sister Organization of ISRO which has also indigenously developed an equally viable infrastructure to put this country on the world map on the space front, owes much to the original vision of Vikram Sarabhai and the sustained and dedicated efforts of Satish Dhawan and his able colleagues.

The AEC and ISRO are examples of what may be called typically 'Indian ethos' developed through truly indigenous efforts. Their infrastructures are sound examples of technological self-reliance which by their very nature could not have materialized without a highly organized structure, which in turn was possible only under a unitary command with direct government control, in company with corresponding structures in developed countries.

State of scientific 'software'

Vital as these organizations are for the health and security of the nation, they are nevertheless a modest fraction of the nation's scientific and technological 'agenda' which span a whole gamut of basic and applied sciences. It is also possible to argue that any worthwhile research efforts in today's state of the art in applied sciences requires a considerable degree of *organized* effort and planning. Moreover, since the infrastructural facilities for such activities cannot be duplicated at will (because of the high cost of installation) their use must also be centralized as far as possible, which explains the emergence of specialized centres like the NSC (with facilities like pelletrons). Similarly, major experimental laboratories for specialized researches in well-defined areas (such as the GMRT) are all 'hard science' activities which necessarily require a high degree of organized and dedicated effort.

The main concern in this review, on the other hand, is not so much an indigenous infrastructure of scientific 'hardware' (AEC, ISRO, etc), crucial as it is for the Nation, as is the development of an indigenous software in the basic sciences for which the traditional breeding ground has been the university system in any country, developed or developing. It is here that one does *not* need any high degree of organization which may well be counterproductive! On the other hand, its relatively informal atmosphere provides the best soil for a gradual and spontaneous evolution of a healthy scientific culture which can be truly claimed as our own. Unfortunately in this field, our performance has been dismal for, not to speak of generating an 'indigenous' flavour, as in most successful nations, it has not even produced a well-defined 'composite' flavour after fifty years of independence, and the blame for this must be laid at the door of our educational policy which simply failed to take off, and whose worst victims have been the university system. This in turn has affected all sectors of the scientific and industrial base, in both basic and applied sciences, as a result of the 'poor software' produced at source! There is no dearth of quantity – about 4000 PhDs per annum – but this itself is more of a liability than an asset, in the absence of a 'sense of direction'. As found in a recent survey[9], this community is beset with a sense of frustration due to loss of self-esteem, lack of intellectual stimulation due to erosion of knowledge base, and a general sense of inadequacy with the environment. All these (and more) are the bane of the typical average scientist of today who has been variously depicted as a 'holy cow' and a 'fall guy' with a poor image of himself, whose labour is unrecognized and his achievements ignored. An analogous picture is also discernable in the general public's perception[10].

As against this plight of the average scientific worker, the scientific culture of the higher echelons among them presents a different kind of scenario, as succinctly described by P. K. Das, Editor, *Society for Scientific Values*, in some recent issues of its bimonthly newsletter[12,13]. In a candid though satirical exposition, the distinguished meteorologist[12] expresses the view that Indian Science is badly afflicted with the so-called 'publicity disease' among its senior and 'high profile' practitioners. In his view, this disease has considerably eroded the real scientific base of the country by sending the wrong signals to the upcoming scientists. Without going into specific examples (several of which are of recent origin), it is perhaps fair to comment that there is nothing intrinsically wrong in wanting quick publicity – this phenomenon is quite universal the world over, provided the claim for priority is endorsed by the scientific community of the worker through a journal publication and/or exposure through a conference presentation. Indeed it is precisely in such situations that the advantages of a 'national ethos' become quite crucial since the projection of such results before the wider world becomes the primary responsibility and initiative of our own scientific community itself. However, if the work is of a secondary nature, which is more often the case, he must accept the initial disadvantage of 'missing the primary bus' and forgo any claims to instant glamour/kudos before doing a more substantial (yet secondary)

investigation, which usually is a time-consuming process with little chance of instant success. Indeed it is up to the senior members of the scientific community to ensure such conformity to the accepted norms of a mature scientific society so as not to send the wrong signals to the younger members of the community as to the correct scientific perspective.

P. K. Das's second article[13] discusses the 'glamour'-oriented high profile behaviour of the higher echelons of our scientific community at the taxpayer's expense, such as (i) travelling from airport to airport to inaugurate symposia in five-star hotels (with speeches written by someone else); (ii) running down the work of others in their regional networks: (iii) and even indulging in subtle forms of plagiarism. An intriguing form of the last type shows up as 'an unduly servile attitude towards scientists from countries other than one's own'–quote from ref. 13. On the other hand, from the higher ups one often hears the oft repeated phrases like 'Scientific temper' and 'Excellence in science', as if to bring out rather tellingly a sort of inverse relationship between preaching and practising, much like two canonically conjugate (complementary?) variables in physics! One hardly hears of such phrases in the western societies where the 'scientific temper' is a routine matter for practice, not for publicity. Why is it the other way round in this country of ours? Otherwise it is again difficult to understand why a 'John Maddox' has to be periodically imported into this country every ten years (roughly coinciding with the Golden and Diamond Jubilees of our premier Academies) just to give a pat on the back of our 'scientists', or a 'certificate of merit' through the columns of *Nature*. In no other country with a mature tradition of science would such self-appointed angels be allowed to exercise such uncalled-for patronage. On the other hand, it is surprising how much can still be done in this very country by purely Indian efforts, if only the right kind of encouragement were provided.

Summary and outlook

This somewhat repetitive analysis was designed for a panoramic view as to the nature of the disease which the scientific community suffers from today, one that is intimately linked with (i) the educational system which simply failed to inculcate the proper sense of values in the initial stages of development, and (ii) with the work culture of the scientific and technological institutions at the later stages of development of the scientific worker. On both fronts he finds himself completely lost in today's scientific labyrinth which seems to be going in tandem with the political atmosphere in the country. The scientific community is merely a part of our bigger society which has been witnessing a steady deterioration of our basic value system – a sort of 'systems failure' from which no part of the society can remain immune. The only difference is that greater maturity and self-restraint would have been expected of the upper echelons of the scientific community which after all represents the 'cream' of the society. One hopes that all is not lost yet, for there are definite indications of renaissance in the major disciplines of science through the coming up of a new generation of workers with the right credentials, with good prospects of their numbers exceeding the critical size, so as to serve as the Conscience of the Scientific community of the future, as well as a watchdog against the kind of aberrations that have been witnessed during the 'transition decades' since Independence.

1. *Quality of Scientific Journals in India*, INSA, April 1993.
2. *University Education in Science*, Indian Acad. Sci., Dec. 1994.
3. Srivastava, P. N., Presidential Address, I. S. C. A., 1993.
4. *Challenging Areas in Physical Sciences*, DST, July, 1990.
5. *Accountability in Scientific Research*, SSV, New Delhi, 1995.
6. Abdus Salam, in a private conversation with the author, 1970.
7. The full text of Nehru's statement appears on the panelled wall of UGC's Conference room no. 114 in Delhi.
8. Sundaram, C. V., *Curr. Sci.*, 1993, **64**, 210.
9. Nagpaul, P. S., NISTADS report Characteristics of Scientific Potential in Academia, 1994.
10. Sridhara, S., What ails science, *Indian Express*, Bangalore, February 11, 1995.
11. Mahajan, S. M. and Sudarshan, E. C. G., *Curr. Sci.*, 1995, **68**, 1260.
12. Das, P. K., On the art of jumping on a bandwagon, *Soc. Sci. Values Newsletter*, 1994, vol. 2, no 4.
13. Das, P. K., On the art of doing nothing, *Soc. Sci. Values Newsletter*, 1995, vol. 3, no 1.

ACKNOWLEDGEMENT. I am grateful to Dr Raja Ramanna, Director, NIAS, for the warm hospitality of the Institute.

A. N. Mitra is in the National Institute of Advanced Studies, Bangalore 560 012, India.

Ethephon enhances karaya gum yield and wound healing response: A preliminary report

Gum karaya (also known as gum kadaya) is the whitish dry exudate from *Sterculia urens* (Sterculiaceae). It is one of the least soluble gums used for many industries such as petroleum and gas, textile, paper and pulp, leather and allied products, ammunition and explosives, electrical appliances, adhesives, confectionery, medicine, pharamaceuticals and cosmetics[1]. Karaya is also known by the name Indian tragacanth as it resembles gum tragacanth produced by *Astragalus* sp[2]. There is high demand for export of gum karaya from India. Yield of this gum has shown a phenomenal decrease from 6838 MT in 1975–76 to 461.3 MT in 1990–91. During this period the price has shot up from Rs. 7491 per MT to Rs. 83, 361 (ref. 1).

The tree has a thick bark (2–3.5 cm) which appears whitish-pink on the outside. Lysigenously formed gum ducts are present in the pith and cortex of the young stem[3], but they are absent in the secondary phloem and wood. However, ducts are formed in the bark and wood in mature trees as a result of injury[4].

Commercial tapping of karaya is done by blazing, peeling or by making deep cuts in the base of the bole with an axe. These methods are wasteful and injurious to the trees, often leading to their death. On account of crude tapping methods, and over-exploitation, the population of karaya trees has markedly declined. In the absence of cultivation of the trees in regular plantations, there is grave concern about the loss of wild germplasm of *S. urens*. Presently, the Governments of Madhya Pradesh, Rajasthan and Uttar Pradesh have imposed a ban on tapping and collection of the gum karaya to allow recovery and regeneration of the trees.

As gum karaya is vital for tribal economy and as its trade value is substantial, there is a pressing need to develop a scientific and sustainable tapping method to increase the yield and to ensure the survival of the tapped trees. Here we present a simple and safe technique for tapping, with a substantial increase in the yield.

Seven trees of various sizes and ages growing in a forest near Ghatti Village along National Highway No.3 in MP (about 25 km from Gwalior), India, were selected for experimentation. Using a hand drill, 10 holes each of 5 mm diameter and 2–3.5 cm depth (based on the thickness of the bark) were bored on each tree around the stem at equal distance, about 1 m above the ground level. The holes were angled towards the base of the tree to prevent backflow of the introduced solutions. Five consecutive holes were treated with ethephon (2-chloroethyl phosphonic acid) and the other five (opposite to the treated holes) were maintained as distilled water controls. One ml of ethephon solution (containing 190/285/390 mg of the active compound per ml) or 1 ml of distilled water was dispensed into each hole.

Table 1. Yield of gum karaya from control and ethephon-treated trees

Distilled water control	Ethephon-treated	
Total amount of gum collected per tree (g)	Amount of ethephon applied (mg)	Total amount of gum collected per tree (g)
a	390	386.00
12.50	390	185.00
9.05	190	433.00
a	190	48.00
8.75	285	65.00
a	285	211.00
15.00	285	259.00
11.30 mean		226.71 mean

a. Gum was not noticed; probably the lump had fallen off before collection.

Gum produced in the control and treated holes was collected after 30 and 45 days of treatment and the data were pooled. The ethephon treatment resulted in an average increase of 20 times more gum than the control (Table 1). The total harvest from seven ethephon-treated trees tapped once was about 1.5 kg of high-quality gum. There was a marked difference in the yield among individual trees, presumably due to heterozygosity.

Notably the control holes were completely dry after 30 days, while the treated ones continued to secrete the gum. Histological examination showed that large gum ducts or cavities had been formed in the secondary phloem and wood, covering a radius of 2–3 cm around each hole in ethephon-treated trees. By 45 days, a thick wound tissue had developed at the injured regions which nearly replaced the damaged tissues.

Ethephon is conventionally used to stimulate latex flow in para rubber (*Hevea brasiliensis*), resin enhancement in pines[5], gum exudation in *Prunus*[6], *Acacia senegal*[7], *Anogiessus latifolia*[8], *Azadirachta indica*[9,10] and increase the production of kino in *Eucalyptus*[11], gum-resinosis in *Commiphora wightii*[12] and *Mangifera indica*[13]. Ethephon is a nontoxic, environment-friendly, inexpensive and easily available plant growth regulator (PGR) manufactured in India and used extensively in agriculture and horticulture[14–16]. Safe tapping and the use of ethephon for improving the yield would ensure sustainable supply of gum karaya.

We are conscious that the sample size on which our observations are based is small. Further detailed studies are being carried out using larger number of trees from separate populations in different seasons to allow statistical assessment of the efficacy of this method. A recent trial with 17 trees has shown that gum yield is high with ethephons treatment even after 10 days. Efforts are being made to standardize the protocol, optimize the yield, and determine the long-term effects on the trees. Yields need to be compared with those from the traditional tapping methods. A detailed investigation on sustained yield, cost effectiveness, extent of injury, and wound-healing response of the trees is also being carried out to verify this preliminary finding.

1. Gautami, S. and Bhat, R. V., *A Monograph on Gum Karaya*, Silver Prints, Uppal, Hyderabad, 1992.
2. Anonymous, *Useful Plants of India*, PID, CSIR, New Delhi, 1992, p. 601.
3. Shah, J. J. and Setia, R. C., *Phytomorphology*, 1976, **26**, 151–158.
4. Purkayastha, S. K., *Indian For.*, 1959, **85**, 126–130.
5. Wolter, K. E., Proceedings of Light Wood Research Co-ordination Council, Annual Meeting, Florida, USA, 1977.
6. Olien, W. C. and Bukovac, M. J., *Plant Physiol.*, 1982, **70**, 547–555.
7. Bhatt, J. R. and Mohan Ram, H. Y., *Curr. Sci.*, 1990, **59**, 1247–1250.
8. Bhatt, J. R., *Curr. Sci.*, 1987, **56**, 936–940.
9. Nair, M. N. B., Patel, K. R., Shah, J. J. and Pandalai, R. C., *Indian J. Exp. Biol.*, 1980, **18**, 500–503.
10. Nair, M. N. B., Bhatt, J. R. and Shah, J. J., *Indian J. Exp. Biol.*, 1985, **23**, 60–62.
11. Hillis, W. E., *Phytochemistry*, 1975, **14**, 2559–2562.
12. Bhatt, J. R., Nair, M. N. B. and Mohan Ram, H. Y., *Curr. Sci.*, 1989, **58**, 349–357.
13. Bhatt, J. R. and Shah, J. J., *Indian J. Exp. Biol.*, 1985, **23**, 330–339.
14. Morgan, P. W., *Acta Hortic.*, 1973, **34**, 41–54.
15. de Wilde, R. C., *Hortic. Sci.*, 1971, **6**, 364–370.
16. Morgan, P. W., *Bot. Gaz.*, 1980, **141**, 337–346.

ACKNOWLEDGEMENTS. We are grateful to the Ministry of Environment and Forests, Government of India, for financial support. We express our sincere appreciation of the cooperation extended to us by the District Forest Officer, Gwalior Range, to carry out the preliminary work in view of its potential benefit. We thank Mr Ashok Kapur of Indian Cropical Pvt Ltd, New Delhi, for a gift sample of ethephon and Mr V. G. Sunnichan for his technical assistance.

M. N. B. NAIR
K. R. SHIVANNA
H. Y. MOHAN RAM

Department of Botany,
University of Delhi,
Delhi 110 007, India

Occurrence of *Botryococcus* from the Lower Cretaceous rocks of Cauvery Basin, Tamil Nadu

Botryococcus, an oil-producing colonial green alga, is generally known to occur in Tertiary and Quaternary rocks of India. It is also reported from Permian sediments of India.

In course of drilling for groundwater in the area, about 1.2 km west and 1.3 km north of Kattarambakam village of Vannur taluk, South Arcot district, Tamil Nadu (Figure 1), carbonaceous shale, sandstone, coal and clay were encountered. The palynological studies of the carbonaceous shale samples revealed the presence of abundant green alga *Botryococcus*[1].

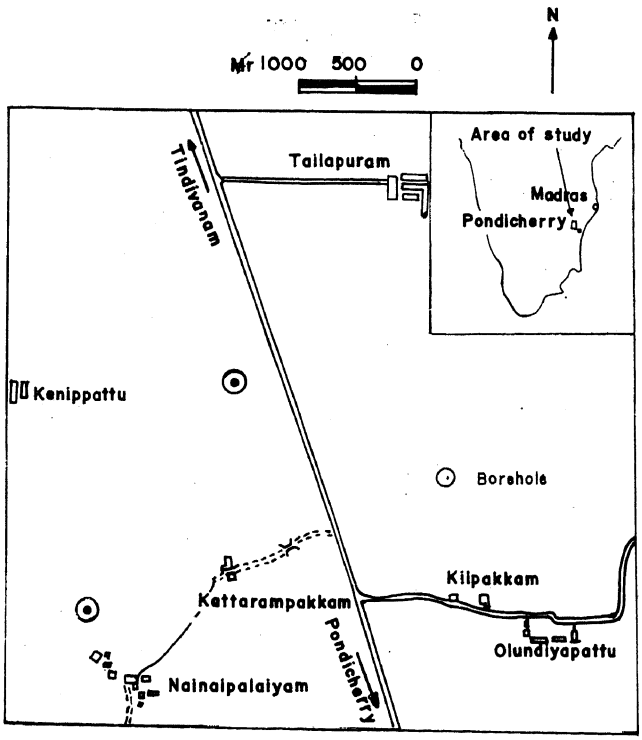


Figure 1. Location map of the study area.

Botryococcus is well preserved, honey yellow to brownish yellow in colour and occurs either as solitary autospores or bodies of two to four chambered cups (Figure 2) or as multicup aggregates. The colonies are mostly botryoidal (Figure 2) or spheroidal in shape and range from 120 to 160 μm in diameter. The diameter of the two- or four-chambered cups ranges from 16 to 28 μm (Figure 2). Further, it is also noticed that some of the colonies appear as irregular lumps of varying dimensions (Figure 2). In the present assemblage the thimble (inner part of cup) invariably remained empty.

Botryococcus is know to thrive in fresh to brackish shallow waters with undisturbed conditions in the area of relatively low rainfall[2]. The excellent state of preservation of all the developmental stages (*i.e.* autospores, two-to-four-chambered cups and multicup aggregates) is indicative of rapid burial of this algal bloom in quiet shallow oxygenated waters without much transportation[2,3].

The qualitative analysis of the associated miofloral assemblage revealed the presence of characteristic palynomorphs referable to *Aequitriradites* sp., *Microcachryidites* sp., *Dictyosporites speciosa*, *Cicatricosisporites australiensis*, *Appendicisporites tricornitatus*, *Neoraistrikia* sp., *Cooksonites* sp., *Gothanipollis*

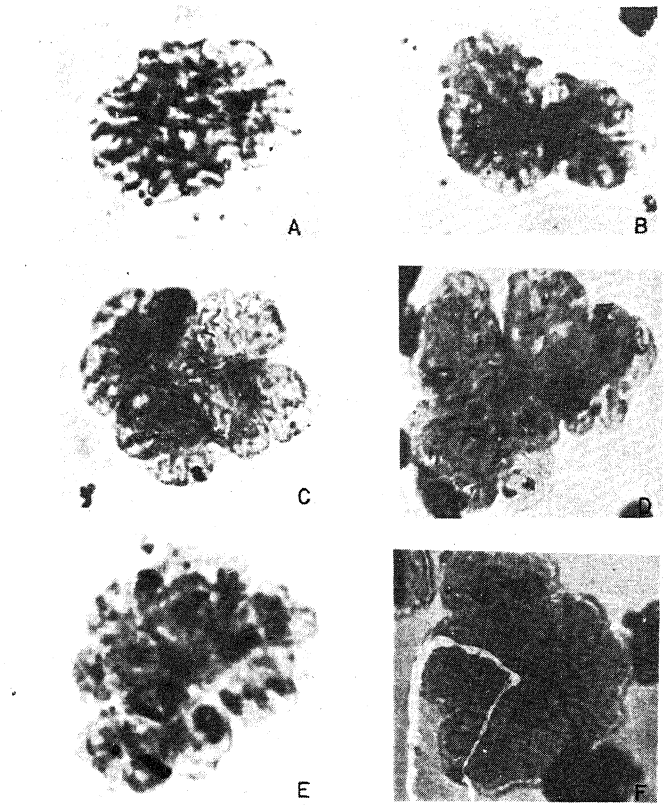


Figure 2. *Botryococcus* showing different stages of preservation: Solitary autospore (× 960); Two-chambered cup (× 960); Four-chambered cup (× 900); Botryoidal nature (× 450); Multicup aggregate (× 900); Abraded form (× 400).

sp., *Crybelosporites stylosus*, *Triporoletes reticulatus* and *T. radiates* (common occurrence) show a close resemblance with the *Triporoletes reticulatus* zone of this *Botryococcus*-bearing palynoassemblage from Kattarambakam.

Most of the fossil records of *Botryococcus* in India are from Eocene[5–7], Miocene[6–9] and Pleistocene sediments[10]. Occurrence of this alga is also reported from Permian rocks of India[11–14]. The present record of *Botryococcus* from the Lower Cretaceous (Late Albian) sediments from Kattarambakam is significant for it is the first record of this alga from Cretaceous of India.

1. Patil, R. S., Rama Rao, M., Gururaja, M. N., Kumaraguru, P. and Kaniskan, B., *GSI News*, 1993, **11**(1), 10.
2. Dorthy, G. O., *Rev. Palaeobot. Palynol.*, 1992, **71**, 1–15.
3. Fristch, F. E., *The Structure and Reproduction of Algae*, Cambridge Univ. Press, 1977, pp. 476–478.
4. Venkatachala, B. S., Rawat, M. S., Sharma, K. D. and Jain, A. R., Proceedings of the IV International Palynological Conference, Lucknow, 1976, pp. 396–402.
5. Lahiri, A., *Proc. Natl. Acad. Sci.*, 1994, 14.
6. Rao, S. R. N. and Mishra, S. S., *Curr. Sci.*, 1949, **18**, 380.
7. Vimal, K. P., *Curr. Sci.*, 1953, **22**, 375–376.
8. Chandra, A., *Curr. Sci.*, 1963, **33**, 214.
9. Ramanujam, C. G. K., *Curr. Sci.*, 1966, **35**, 367–368.
10. Chatterjee, N. N. and Chandra, D., *Geol. Min. Met. Soc. India*, 1957, **29**, 187.
11. Sinha, Veena, *Palaeobotanist*, 1969, **17**, 326–337.
12. Bharadwaj, D. C. and Dwivedi, A., *Palaeobotanist*, 1981, **27**, 21.
13. Lele, K. M. and Shukla, M., *Geophytology*, 1980, **10**, 231–238.
14. Srivastava, S. C. *Geophytology*, 1980, **10**, 62–71.

ACKNOWLEDGEMENT. We thank Sri D. C. Das, Geological Survey of India, Hyderabad for encouragement and keen interest in the present investigation.

R. S. PATIL
M. RAMA RAO
M. N. GURURAJA
P. KUMARAGURU*
B. KANISHKAN*

Geological Survey of India,
Regional Palaeontology Laboratory,
Bandlaguda, Hyderabad 500 068, India
**Geological Survey India,*
Inner Ring Road, Guindy,
Madras 600 032,, India

Enechelon faults along West Coast of India and their geological significance

Subsequent to the advent of modern remote sensing technology, mapping and analysis of various lineaments and fractures and their correlation with geophysical and ground geological data have become a matter of great interest to geoscientists. The West Coast of India is one such region which has always attracted the attention of geoscientists and many have studied the region. The significant amongst them are the studies by Jacob and Narayanaswamy[1], Ramachandran[2] and Nair[3]. Jacob and Narayanaswamy[1] have observed that the Western continuity of the Palghat gap extends right up to 9° channel, separating Laccadives and Maldives. Ramachandran[2] has carried out very exhaustive analysis of lineaments by integrating total intensity aeromagnetic, Bouger gravity anomaly and Landsat lineament data sets. On the basis of such an integration, he has identified four sets of lineaments in WNW–ESE to NW–SE, NNW–SSE, N–S to NNE–SSW and NE–SW directions.

Nair[3] has carried out structural interpretation for the entire West Coast of India and brought out certain newer information on the neotectonic activities.

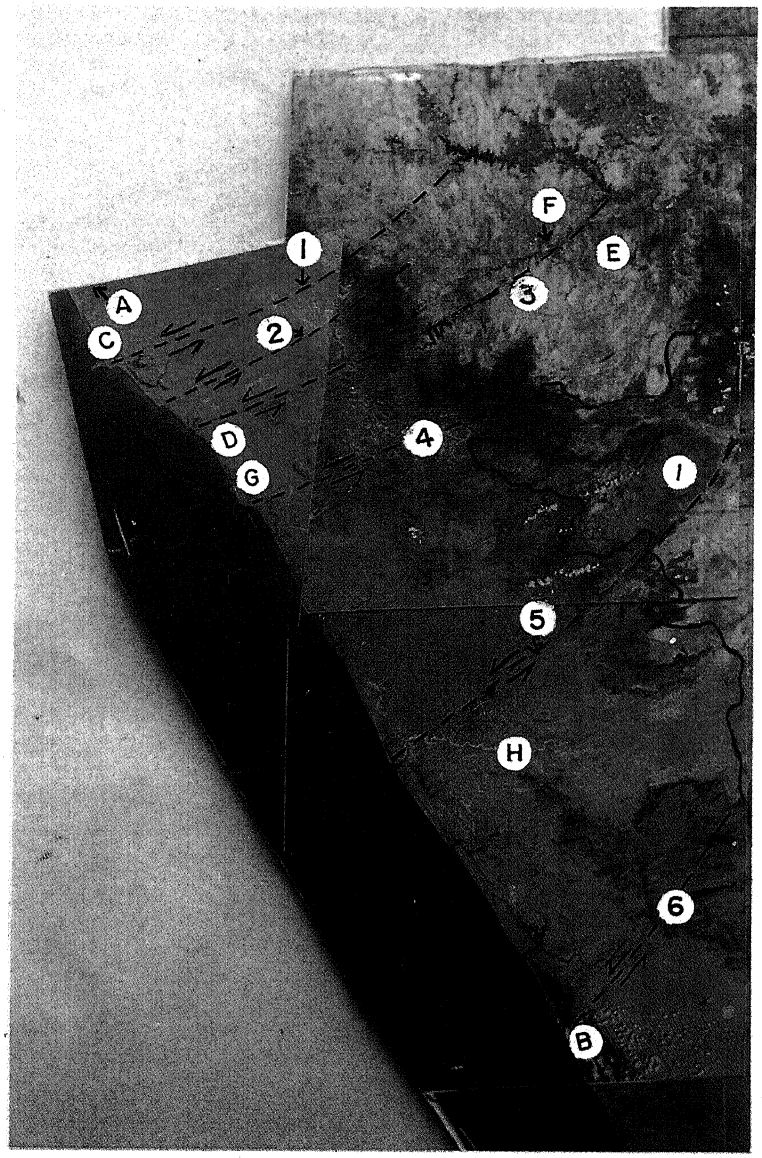


Figure 1. IRS 1A Satellite imagery mosaic of West Coast of Kerala showing Pleistocene faults. 1–6, lineaments/faults; ⊠ Sinistral faults, A, Kasargod; B, Cochin; C, Cannanore; D, Mahe; E, Mysore; F, Kabbani river; G, Pandalayini; H, Ponneni river; I, Ootacamund massif.

In the present study, I have filtered out only the lineaments showing clear shifting of beds in the coastal landforms, which mostly fall in ENE–WSW direction.

The interpretation of IRS IA satellite imagery data for the part of West Coast of Kerala and Karnataka, especially the area falling between Kasargod in the northwest and almost Cochin in the southeast (Figure 1), shows a set of ENE–WSW to NE–SW-trending subparallel lineaments, exhibiting clear sinistral shift of coastal landforms/shorelines in the area. Amongst these lineaments/faults, most of them are continuing even up to 700 to 800 km in the continent with a clear northeasterly swing. They also continue up to Laccadives and Maldives in the southwest with ENE–WSW orientation (Figure 2).

1. The fault no. 1 trends in ENE–WSW direction with clear sinistral drag of beds in the coastal zone and coincides with the Lakshmithirtham lineament of Ramachandran[2].

2. The ENE–WSW trending fault no. 2 of Cannanore region exhibits sinistral drag.

3. The fault no. 3 is again an ENE–WSW-trending linear feature, extending from Mahe in the West Coast to Mysore in Karnataka, with the swing towards northeasterly. This sinistral coastal zone fault controls Kabbani river in the continental part and is described as Kabbani lineament with dextral strikeslip configuration by Katz[4] and Drury and Holt[5].

4. The fault no. 4 shows sinistral shift of 5–6 km in the West Coast in Pandalayini area and merges with WNW–ESE trending Moyyar shear.

5. The fault no. 5 trends in general NE–SW direction from Ponnani river mouth in the West Coast to Elagiri hills in the northeast. This also exhibits minor sinistral shift of beds along Ponnani river mouth. It controls the Bhavani river and delimits the southeastern and southern boundaries of Ootacamund massif and Biligirirangan hills respectively. In its northeastern proximity, it forms the western limit of Javadi hills.

6. The fault no. 6 is trending in NE–SW direction from Cochin in the West Coast to Coimbatore in the northeast where seismicity was reported in 1900 (ref. 6).

Most of the above observed sinistral strikeslip faults are found to extend right up to Laccadives and Maldives in the Arabian sea (Figure 2) and amongst these, the fault nos 5 and 6 are found to form 9° channel, separating Laccadives and Maldives. In addition to the above faults, few more sinistral strikeslip faults were also interpreted (fault nos 8, 9 and 10, Figure 2). All these faults trend in NE–SW direction in the continental part, take a swing towards westsouthwesterly and finally ends in Laccadives and Maldives (Figure 2). Similar to fault no. 5, the fault no. 7 observed south of Alleppey continues as 8° channel in Laccadives, causing similar sinistral shifts of land in Laccadives. All the above discussed ten faults have gradually shifted not only the West Coast of Karnataka and Kerala but also Laccadives and Maldives into an enechelon pattern (Figure 2).

The above discussed six faults (nos 1 to 6, Figure 1) and the other (nos 7 to 10, Figure 2) system of faults have been demonstrated to be Precambrian dextral strikeslip faults in parts of Tamil Nadu and the adjacent states[2,4,5,7].

Such Precambrian NE–SW-trending dextral faults swing towards westsouthwesterly, with sinistral strikeslip configuration in Kerala and Karnataka coast and in Laccadives and Maldives according to the present observations (Figures 1 and 2).

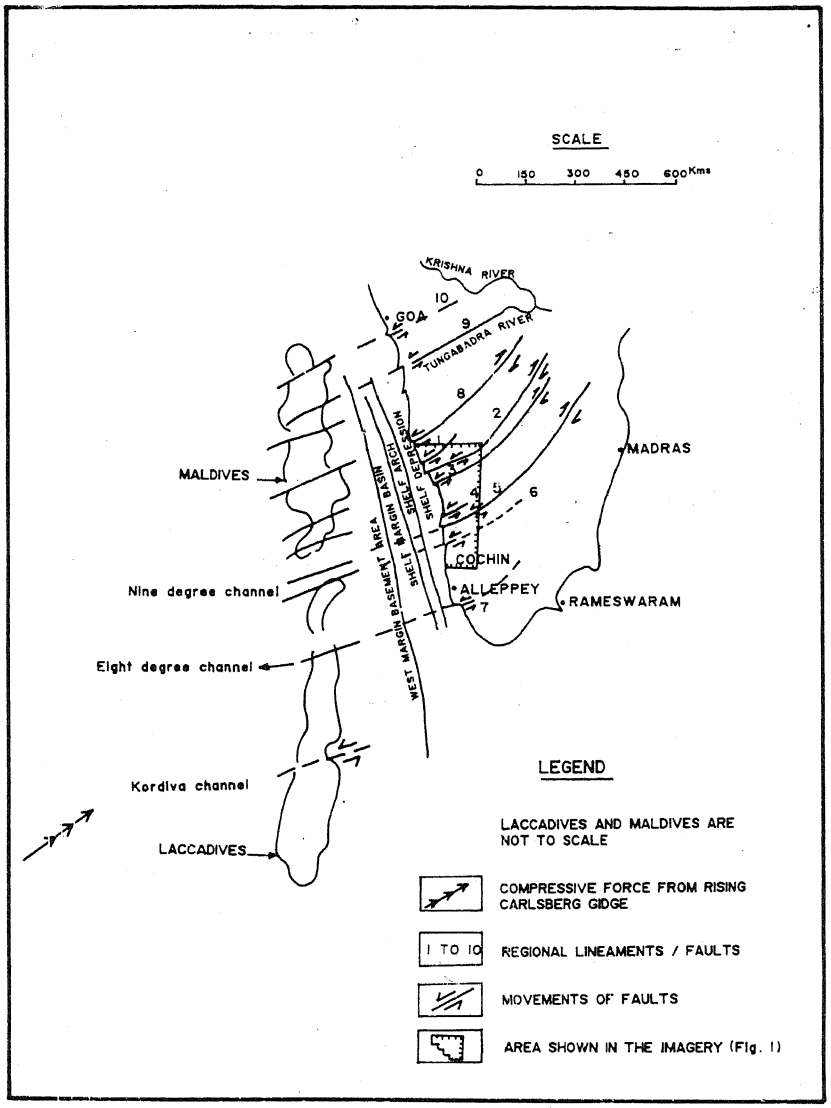


Figure 2. Off-shore and on-shore tectonic features of West Coast.

Nair[8] has also demonstrated clear Holocene earth movements along a few ENE–WSW faults in the coastal plains and also in the surf zone of Kerala coast in the form of lagoons, submerged platform, etc. He also extended few sets of ENE–WSW-trending Holocene faults right up to Laccadives and Maldives.

But, Jacob and Narayanaswamy[1] have extended the Palghat fault up to 9° channel in Laccadives and Maldives. On the contrary, the present study and also the studies by Nair[8] advocate the possibility of extending only the ENE–WSW-trending fault right up to 9° channel and similarly the other faults up to Laccadives and Maldives. Ramasamy and Balaji[7] have once again found that almost all the ENE–WSW-trending faults of Kerala and Karnataka are Pleistocene sinistral faults. Thus, it indicates that all the Precambrian NE–SW-trending regional dextral strikeslip faults of the continental part have been transformed into Pleistocene sinistral faults in the West Coast.

Ramasamy *et al.*[9] have demonstrated an E–W-trending cymatogenic arch between Mangalore in the West Coast and Madras in the East Coast. Subrahmanya[10] too found an upward along Mangalore and Madras alignment.

The integration of such an E–W-trending cymatogenic arch with present set of ENE–WSW to NE–SW-trending sinistral strikeslip faults indicates that these ENE–WSW trending sinistral faults of Kerala are referrable to Left lateral wrenches[11]. And if so, the E–W-trending Mangalore and Madras arch and the sinistral strikeslip faults might have been formed due to the northerly directed compressive force which may be related to the movement of Indian plate towards Tibetan plate.

The possibility for such northerly directed compressive force as a source to Mangalore–Madras arch has also been suggested by Ramasamy *et al.*[9].

An alternate model is also possible to explain these sinistral faults. Lepichon[12] has observed that the Carlsberg ridge is raising at the rate of 1 to 30 cm per year, thereby causing northeasterly directed force towards Indian subcontinent. Radhakrishna[13] has observed NNW–SSE-trending onshore Quaternary fractures in the West Coast of India and attributed the same to an ENE–WSW-trending post-trappean compression.

Ghosh and Zutshi[14] have identified a number of NNW–SSE-trending continental depressions, continental shelf arches and shelf margin basins (Figure 2). All these observations suggest that the West Coast region is being subjected to ENE to northeasterly directed force due to the rise of Carlsberg ridge.

This might have caused not only onshore and offshore NNW–SSE-trending arches, deeps and flexures but also would have transformed the pre-existing Precambrian dextral faults into sinistral ones. However, the morphology and the dynamics of such coastal faults need deeper studies for a better understanding of post-collision tectonics.

1. Jacob, K. and Narayanaswami, S., *Proc. Natl. Inst. Sci. India*, 1954, **20**, 104–108.
2. Ramachandran, T. V., *J. Indian Min.*, 1985, **39**, 17–36.
3. Nair, M. M., *J. Geol. Soc. India*, 1990, **35**, 99–105.
4. Katz, M. B., *J. Geol. Soc. India*, 1978, **19**, 185–203.
5. Drury, S. A. and Holt, R. W., *Tectonophysics*, 1987, **65**, PT1–PT5.
6. Srinivasan, V., *Proceeding Volume of Earth and Atmospheric Sciences of Tamil Nadu Science Congress*, Bharathidasan University, Tiruchirapalli, 1992, pp. 43–52.
7. Ramasamy, SM. and Balaji, S., *Int. J. Remote Sensing*, 1994 (in press).

8. Nair, M. M., *Mem. Geol. Soc. India*, 1987, **29**, 450–458.
9. Ramasamy, Sm., Panchanathan, S. and Palanivel, R., Proceedings of International Geoscience and Remote Sensing Symposium, Michigan, 1987, pp. 1157–1161.
10. Subrahmanya, K. R., *Curr. Sci.*, 1994, **67**, 527–530.
11. Moody, J. D. and Hill, M., *J. Bull. Geol. Soc. Am.*, 1956, **67**, 1207–1246.
12. Lepichon, X., *J. Geophys. Res.*, 1968, **73**, 3661–3697.
13. Radhakrishna, B. P., in Proceedings of the Seminar on Geomorphological Studies in India, Univ. of Saugar, Sagar, 1967.
14. Ghosh, B. N. and Zutshi, P. L., *Geol. Surv. Ind. Spl. Publ. No. 24*, 1987, pp. 309–318.

Sm. RAMASAMY

Centre for Remote Sensing, School of Earth Sciences, Bharathidasan University, Tiruchirapalli 620 023, India

Interpretation of nucleotide/protein sequence data: Some pitfalls

Identification of unknown genes and their products by computer search have received high priority among molecular biologists. However, since biologists are not always computer analysts at the same time, there always exists a finite probability of making errors in analysing sequences and in interpretation of data which can easily be avoided with some advice from computer analysts. We would like to highlight some of the precautionary measures that need to be taken to avoid the inherent pitfalls in sequence data and their proper interpretations.

The worst and a very common source of errors in DNA sequences is the contamination of vector sequences arising during cloning and subcloning of DNA fragments. In 20,000 sequences in the GenBank 63 alone, more than 50 instances of cloning vector contamination have been detected[1]. In a smaller number of cases the anomalous sequences might have arisen during editing, but in a majority of the cases, large blocks of vector sequences contaminated the actual sequence. Incorporation of anomalous sequences, particularly in the coding regions, not only hinders the recognition of low-level homologies and consensus sequences, but can also exhibit false similarities with other sequences, leading to erroneous conclusions. An example of such an error is in the nucleotide sequence of the *recA* gene of *Vibrio cholerae*[2] published from our laboratory. While careful sequencing of both strands and comparison of the sequence-based restriction maps with the map of the cloned DNA fragments might eliminate the problem, vector contamination can be easily and quickly detected and removed using computer programs that are now available[1]. The contamination can also be removed by comparing the sequence with the dataset of vector sequences available at the GenBank.

'Frameshift' is another frequently occurring error in DNA sequences[3,4]. This happens when a base is either missed or added during reading of sequencing gels. Posfai and Roberts[5] have detected many such errors in EMBL (release 24) and GenBank (release 56) using the program DETECT, which examines alternative reading frames from related proteins. The program BLASTX[4,6] based on BLAST[7] algorithm can also predict frameshift errors by comparing the translated nucleotide sequences from all six reading frames with a protein database. When the similarity to a protein switches from one frame to another in the same strand of the query sequence, there is definitely a frameshift error in the sequence[4]. However, with the simultaneous translation and alignment algorithms, proteins with $>30\%$ sequence identity can be reliably recognized even in the presence of 1% frameshifting error rates and 5% base substitution rates[8].

In nucleotide sequence analysis, unknown complete or partial open reading frames (ORFs) are often encountered along with the sequence of the gene of interest[4,6]. These ORFs, overlooked in most cases by the investigators, might represent useful functions. Two computational approaches are now used either independently or in combination to determine whether the overlooked putative ORFs are indeed bonafide genes[3,4,6]. The first is an intrinsic approach, which distinguishes the coding regions from the non-coding ones on the basis of the statistical analysis of some parameters of the sequence without referral to any other sequence[4,6,9,10]. The software that has been extensively used for this purpose in human genome sequence analysis[2] is GRAIL[11], which utilizes the sensor-neural network approach to evaluate the intrinsic properties of the DNA sequence like frame bias, dinucleotide fractal occurrence, etc. This approach can also be used in the analysis of small genome after proper modification. Using the GenMark method based on phased Markov Chain model, Borodovsky *et al.* predicted expressed ORFs in the unannoted regions of *Escherichia coli* genomic DNA from EcoSeq6 database[12]. The second approach for predicting genes is extrinsic and involves comparison of the putative deduced amino-acid sequence with protein sequence databases and searching for motifs[4,6]. If the deduced amino-acid sequence of the putative ORF shows 'significant' similarity to one or more proteins in the database, it is almost certain that the putative ORF represents a bonafide gene. It may be pseudo- or cryptic gene, but is definitely not a part of non-coding region[4]. Using these two approaches, a large number of new genes in the unannoted regions of *E. coli* genome have been identified and searches for genes that have escaped detection so far in several other organisms are in progress.

To search for homology of a newly generated protein sequence with other sequences, i.e. databank one or the other alignment programs are used[7, 13–15]. If

after suitable gaping more than 25% residues of two protein sequences longer than 100 residues are identical, they must be evolutionarily and/or functionally related[16]. If the identity is between 15 and 25%, the sequences are considered to be marginally similar and in the language of computational biology, this range is known as 'twilight zone'[16]. Although this criterion of assessing 'similarity' between two sequences is acceptable, a more rigorous approach is provided by the BLAST program[7] where a probability factor (p) is assigned to each alignment, which decides whether the observed similarity between two sequences could have been obtained by chance[4]. Usually $p < 10^{-4}$ suggests genuine homology between two sequences and $10^{-4} < p < 10^{-1}$ constitutes the 'twilight zone'[4]. When the similarity between the two sequences falls in the twilight zone, some caution should be taken to decide whether the sequences are genuinely related or the twilight alignment is an artifact. Ignoring alignments with $p > 10^{-4}$ might eliminate the possibility of detecting false relatedness, but such stringent restrictions might hinder detection of weak but genuine relatedness between two sequences. What is recommended in these cases is to check whether the alignment in the twilight zone is conserved in a protein family. The software MACAW[17] can detect subtle similarities between protein sequences.

Protein sequences which have 'biased amino acid composition' often generate erroneous conclusions regarding its relatedness with other proteins with similar type of bias. For example, heat-shock proteins have clusters of charged amino-acid residues[18], the human transcription factors are rich in glutamine[19], zinc finger proteins and homeo-box regions of DNA-binding proteins have bias towards positively charged amino-acid residues[20]. If a new sequence has any of these characters, it will show similarity in the region of bias even when the proteins are functionally not related. This problem can be solved by filtering out the regions of bias before searching for similarity with other proteins. The program SEG can filter out the biased or 'low-complexity' regions from a sequence[6].

Computers function objectively and in computer searches there is no scope for prejudice that is often associated with experimental studies. This advantage of computational methods, the explosive growth in sequence databases and the availability of a wide variety of analytical software packages have, in recent years, encouraged molecular biologists to undertake computational analysis of nucleotide and protein sequences to derive a variety of information. To exploit fully the sequence data it is, however, necessary to assimilate new approaches of computational methods and this can only be realized through users having strong background not only in molecular biology but also in computers.

1. Lamperti, E. D., Kittelberger, J. M., Smith, T. F. and Villa-Komaroff, L., *Nucleic Acids Res.*, 1992, **20**, 2741–2747.
2. Ghosh, S. K., Biswas, S. K., Paul, K. and Das, J., *Nucleic Acids Res.*, 1992, **20**, 372.
3. Doolittle, R. F., *Curr. Op. Biotech.*, 1994, **5**, 24–28.
4. Koonin, E. V. and Rudd, K. E., *Trend Biochem. Sci.*, 1994, **19**, 309–313.
5. Posfai, J. and Roberts, R. J., *Proc. Natl. Acad. Sci. USA*, 1992, **89**, 4698–4702.
6. Borodovsky, M., Rudd, K. E. and Koonin, E. V., *Nucleic Acids Res.*, 1994, **22**, 4756–4767.
7. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J., *J. Mol. Biol.*, 1990, **215**, 403–410.
8. States, D. J. and Botstein, D., *Proc. Natl. Acad. Sci. USA*, 1991, **88**, 5518–5522.
9. Staden, R., *Methods Enzymol.*, 1990, **183**, 163–180.
10. Bougueleret, L., Tekaia, F., Sauvaget, I. and Claverie, J. M., *Nucleic Acids Res.*, 1988, **16**, 1729–1738.
11. Uberbacher, E. C. and Mural, R. J., *Proc. Natl. Acad. Sci. USA*, 1991, **88**, 11261–11265.
12. Rudd, K. E., *ASM News*, 1993, **59**, 335–341.
13. Lipman, D. J. and Pearson, W. R., *Science*, 1985, **227**, 1435–1441.
14. Smith, T. F. and Waterman, M. S., *J. Mol. Biol.*, 1981, **147**, 195–197.
15. Feng, D. F., Johnson, M. S. and Doolittle, R. F., *J. Mol. Biol.*, 1985, **21**, 112–125.
16. Doolittle, R. F., in *Of URFs and ORFs*, Univ. Science Books, Mill Valley, 1987.
17. Tatusov, R. L., Altschul, S. F. and Koonin, E. V., *Proc. Natl. Acad. Sci. USA*, 1994, **91**, 12091–12095.
18. Karlin, S., Blaisdell, B. E. and Brendel, V., *Methods Enzymol.*, 1990, **183**, 388–402.
19. Courey, A. J. and Tijan, R., *Cell*, 1988, **55**, 887–898.
20. Brendel, V. and Karlin, S., *Proc. Natl. Acad. Sci. USA*, 1989, **86**, 5698–5702.

CHITRA DUTTA
JYOTIRMOY DAS

Biophysics Division,
Indian Institute of Chemical Biology,
Calcutta 700 032, India

Fifty years of the exact solution of the two-dimensional Ising model by Onsager

Somendra M. Bhattacharjee and Avinash Khare

The exact solution of the two-dimensional Ising model by Onsager in 1944 represents one of the landmarks in theoretical physics. On the occasion of the fifty years of the exact solution, we give a historical review of this model. After briefly discussing the exact solution by Onsager, we point out some of the recent developments in this field. The exact solution by Onsager has inspired several developments in various other fields. Some of these are also briefly mentioned.

1. Introduction

It is usually said that every theoretical physicist must know (ideally must be taught) the landmarks of theoretical physics. Some of these are: g-2 calculation in quantum electrodynamics, Onsager's exact solution of the two-dimensional Ising model etc. It is our privilege to briefly discuss Onsager's exact solution of the Ising model and its influence in the subsequent developments of statistical mechanics[1]. To appreciate why it is regarded as one of the landmarks of theoretical physics, it is worth remembering that for a long time it remained the first and the only (mathematically rigorously) exactly solvable model exhibiting phase transition. Its discovery completely changed the course of development in statistical mechanics and also other areas of physics. Before Onsager's exact solution, it was not clear if the formalism of statistical mechanics could handle phase transition. The solution established beyond doubt that phase transitions appear as singularities in the thermodynamic functions and these functions need not have simple discontinuities as hypothesized by Ehrenfest before. Furthermore, of all the systems in statistical mechanics on which exact calculations have been performed, the 2-dimensional Ising model is not only the most thoroughly investigated but it also the most profound. Its significance was instantly recognized. In this context we would like to recall the letter written by Pauli to Casimir immediately after the World War II. Casimir in his letter had expressed his concern about being cut off for so long from theoretical physics of allied countries. Pauli in his reply said 'nothing much of interest has happened except for Onsager's exact solution of the two-dimensional Ising model'[2].

The plan of the article is as follows: In §2 we first define the Ising model and give an historical account of the various approximate solutions developed leading to the exact solution of Onsager[3]. We shall also briefly mention the relevance of the model. A short sketch of the life and works of Onsager is given in §3. In §4 we review the exact solution of Onsager[4] and its importance. Some of the open, unsolved problems are also mentioned here. It must be made clear here that our aim is not really to discuss the exact solution since there are several places where an excellent account has already been given. In §5 we discuss the various developments in this field which have been directly inspired or influenced by Onsager's work.

Somendra M. Bhattacharjee and Avinash Khare are at the Institute of Physics, Sachivalaya Marg, Bhubaneswar 751 005, India.

2. The model and its brief history

The Ising model was in fact first written down not by Ising but by his thesis supervisor W. Lenz[5] in 1920 who was then working in Rostalk University in Germany. It is somewhat unfortunate that the physics community has given him no credit for this work. This is perhaps because even though Lenz introduced the model, he did not do any calculations using this model. It is perhaps worth recalling that Lenz is instead well-known for his work on Runge–Lenz vector in the context of the accidental degeneracy in the hydrogen atom problem. Lenz moved to Hamburg as professor in 1921. One of his first Ph D students was Ernest Ising who was born on May 10, 1900. In late 1922, Lenz asked Ising to study his model and the phenomena of ferromagnetism. Ising studied the model and found its exact solution in one dimension and showed that there is no phase transition from para to ferromagnetism[6]. Since then the model is known as the Ising model. Let us now briefly state the Ising model.

The model

It is a lattice model. Consider a d-dimensional periodic lattice (d = 1,2,3,...) having an array of N fixed points. The lattice may be of any type. For example, one could

have 3-dimensional cubic or hexagonal lattice. However, we shall mostly considering the two-dimensional square lattice since it is for this that Onsager obtained his exact solution. With each lattice site is associated a spin variable S_i ($i = 1,2,..., N$) which is a number being +1 or −1. If it is 1 or −1 we call it spin up or spin down. Clearly, a given set of numbers $\{S_i\}$ specifies a configuration of the whole system and that there are in all 2^N different configurations. The energy of the system in a given configuration S_i is defined to be

$$E\{S_i\} = -\sum_{<ij>} J_{ij} S_i S_j - B\sum_{i=1}^{N} S_i, \qquad (1)$$

where B is the external magnetic field. Usually one assumes that the interaction is isotropic (i.e. $J_{ij} = J > 0$) and that there is only nearest-neighbour interaction. For example, for a 2-dimensional square lattice, the number of nearest neighbours is 4. Given the energy of the system, the goal is to compute the canonical partition function ($\beta = (kT)^{-1}$, where k is the Boltzmann constant and T is the temperature)

$$Q(B,T) = \sum_{S_1}\sum_{S_2}\cdots\sum_{S_N} e^{-\beta E\{S_i\}}, \qquad (2)$$

and hence the thermodynamic properties, and to see if the model exhibits and phase transition at a finite nonzero temperature. In other words, in magnetic systems, each molecule has a spin that can orient up or down relative to the direction of the applied external magnetic field B. The question one would like to ask is if this model could lead to spontaneous magnetization i.e. if below a certain critical temperature T_c, $M(B = 0, T)$ which is essentially same as the order parameter $\bar{L}(B = 0,T)$, is nonzero or not. Here the magnetization $M(B,T)$ is related to the partition function by

$$M(B,T) = +\frac{\partial}{\partial B}[\ln Q(B,T)].$$

It is worth pointing out here that the Ising model equally well represents a model for (i) binary alloys and (ii) lattice gas. Binary alloys are mixtures of two types of molecules (for example β-brass has a cubic structure made out of Zn and Cu atoms) and the question is if below a certain temperature T_c, there is a phase transition with atoms of the same type clustering together. On the other hand, in lattice gas models, one considers a mixture of molecules and holes (i.e. empty spaces) on a lattice and the question is if below T_c, there is a condensation of molecules into one region of space and holes in the rest of the lattice? Yang and Lee[7] discussed this model and gave a detailed mapping between it and the Ising model. It may be noted here that the lattice gas models are of relevance in the context of gas–liquid and liquid–solid transitions. We thus see that the study of the Ising model is of relevance in a number of phenomena. Infact for $J < 0$, this model could very well represent (I) antiferromagnetic ordering, (ii) superlattice structure in any alloy, and (iii) a solid-like arrangement of molecules with repulsive forces.

A point in order at this stage. Many people tend to dismiss the Ising model as an over-simplified representation of intermolecular forces. However the point to note is that the essential features of the cooperative phenomena (i.e. long range order), specially near T_c, do not depend on the details of the intermolecular forces but on the mechanism for propagation of long range order and Ising model offers much hope for an accurate study of this mechanism.

Historical developments

As mentioned above, Ising was able to obtain the exact solution of the model in 1-dimension[6] and show that there is no phase transition at $T \neq 0$. He then gave some arguments and erroneously concluded that even in two or three dimensions this model will not exhibit any phase transition. This dissuaded several people from working on this model and prompted Heisenberg[8] to introduce more complicated vector interaction between spins known popularly as the Heisenberg ferromagnetic interaction. It is interesting to note that Ising seriously believed in his conclusions and felt very frustrated that the model has no usefulness and gave up physics research! Being a Jew, he was prosecuted in Germany and was dismissed from his job in 1933. He managed to run away from Germany and lived in Luxemberg from 1939 to 1947. During all this time he had no contact with physics and lost track of future developments about the model. Only in 1947 when he went to USA to teach did he realize that his name has become immortal because of the work done by other people!

Almost a decade after Ising's work, Bragg and Williams[9] studied the model in the mean field approximation. They were inspired by an earlier work of Gorsky[10]. The essential assumption here is that the energy of an individual atom in any configuration is determined by the average degree of order prevailing in the entire system rather than by the fluctuating configurations of the neighbouring atoms. Clearly, this approximation is exact in the limit $d \to \infty$. One of the serious criticisms of this approximation is that it is independent of the dimension d. This approximation predicts a phase transition in the Ising model at finite T for any d which is clearly wrong for $d = 1$ as one knows from Ising's exact calculation.

Soon afterwards, Bethe[11] improved upon the Bragg–Williams approximation by treating somewhat more accurately the interaction between the nearest neighbours. It turns out that this approximation is in fact exact for 1-dimension and hence does not predict phase transition at any finite nonzero T while for 2 and higher dimensions,

it does predict spontaneous magnetization. Soon afterwards, Peierls[12] wrote a paper entitled 'On Ising's model of ferromagnetism' in which he gave a simple argument showing that at sufficiently low temperature, Ising model in 2 or 3 dimensions must exhibit spontaneous magnetization. It turned out much later that his argument involved an incorrect step[13] but nevertheless, the conclusion and the general procedure are correct. This was an important step, and the method is used even now in various situations.

The first exact quantitative result for the 2-dimensional Ising model was obtained by Kramers and Wannier (see ref. 15) when they located the transition temperature by using the symmetry of the two-dimensional lattice to relate the high- and low-T expansion of the partition function. On February 28, 1942 there was a meeting of the New York Academy of Sciences in which Wannier gave a talk on this work. At the end of his talk, Onsager made a remark announcing that he has been able to obtain the exact solution of the 2-dimensional Ising model for a square lattice and without external magnetic field. It is very significant though, that he published his results in a journal only two years later. In fact this was always his style. He could see through the math and get the final result but was careful to publish his results only after he had satisfactorily settled all the mathematical questions.

3. Onsager – His life and works

Lars Onsager was born at Oslo in 1903. Even though he got his degree in chemical engineering in 1925 he devoted most of his time to studying math and physical sciences. In 1931 he wrote two monumental papers in *Physics Review* about reciprocal relations in irreversible processes. Eventually, for this work, he received Nobel Prize in Chemistry in the year 1968. The Nobel citation noted that the publication (two papers in *Physical Review* of length 22 and 15 pages) was one of the smallest ever to be awarded a Nobel prize! In 1933 he was offered the Sterling post-doctoral fellowship at the Chemistry Department of Yale University. But there was one small problem. Onsager had never written his Ph D thesis! He was persuaded to write one which he did on the properties of Mathieu functions. This work is recognized as one of the significant contributions to the subject. In 1934 he became the faculty member of the department. From 1944 till his retirement in 1972 he occupied the J. Willard Gibbs chair of Theoretical Chemistry at Yale University. In 1972, he moved to Miami University where he worked in the Centre for Theoretical Studies till his death in 1976.

During his entire career, Onsager only published about 60 papers but each was an important one published only after he was fully satisfied with its significance and its mathematical rigour. Many a time he announced his results in conferences as a remark after some talk but published the results much later or never did! We will see one such example below. He would call himself a chemist but truly speaking he was one of the last true universalists of this century. For more information see ref. 14.

4. Onsager's exact solution of the 2-D Ising model

To appreciate Onsager's exact solution, it is worth recalling the work of Kramers and Wannier[15,16] before Onsager where they showed that the partition function for this problem can be written as the largest eigenvalue of a certain matrix. Onsager's method was similar to these authors except that he emphasized the abstract properties of rather simple operators rather than their explicit representation by unwidely matrices. It must be said here that his method is highly nontrivial and complicated. Actually his derivation can be easily followed step by step but the overall plan is quite obscure. Since a simplified version as given by Bruria Kaufman[17] has been discussed at a number of places including textbooks we shall merely quote the result. Onsager showed that the canonical partition function $Q(B=0,T)$ in the limit $N \to \infty$ is given by

$$\lim_{N\to\infty} \ln Q(B=0,T) = \ln(2\cosh(2\beta J)) + \frac{1}{2\pi}\int_0^{\pi} d\phi$$

$$\ln\frac{1}{2}\left(1+\sqrt{1-\kappa^2\sin^2\phi}\right), \qquad (4)$$

where $\kappa \equiv 2\sinh(2\beta J)/\cosh^2(2\beta J)$. From here it followed that the specific heat $C(B=0,T)$ defined by

$$C(B,T) = k\beta^2 \frac{\partial^2}{\partial\beta^2}[\ln Q(B,T)], \qquad (5)$$

diverges logarithmically at $T \to T_c$ where T_c is given by

$$\tanh\frac{2J}{kT_c} = \frac{1}{\sqrt{2}} \Rightarrow \frac{kT_c}{J} = 2.269185 \qquad (6)$$

For comparison, kT_c/J is predicted to be 4 and 2.88 in the mean field and the Bethe approximations respectively.

Over the years, several alternative simplified methods have been given but all of them are quite involved and lengthy. Bruria Kaufman[17] gave a simplified proof which is based on spinorial representation of the rotation group and it is this proof which is essentially given in K. Huang's book on statistical mechanics. Further, using Onsager's proof, the exact partition function and hence other properties of several other 2-dimensional lattices have been deduced. Finally, Onsager and Bruria Kaufman[18] calculated spin correlation functions

Further developments

To justify calling the phenomenon at $T = T_c$ a phase transition, one has to examine the long range order i.e.

spontaneous magnetization and show that it is nonzero. This is a very difficult calculation since it has to be done with $B \neq 0$ (but could be small) at the beginning of the calculation and finally putting it equal to 0 at the end of the calculation. On August 23, 1948 there was a talk by Tisza at Cornell University. At the end of the talk, Onsager walked up to the blackboard and coolly announced that he and Bruria Kaufman had solved this problem and they indeed found that the long range order is nonzero. Infact he even wrote down the formula on the blackboard. He repeated his comments at the first postwar IUPAP conference on Statistical Mechanics at Florence in 1949 after a talk by Rushbrook. However, Kaufman and Onsager never published their calculations; they only appeared as a discussion remark[19]. In print the first full calculation was in fact published by Yang[20]. Why did Onsager not publish his results? Years later, only in 1969 did he give the reason. In computing the long range order, Onsager was led to a general consideration of Toeplitz matrices but he did not know how to fill out holes in the maths – by the time he did, mathematicians were already there! What Onsager and Yang had obtained was that as T goes to T_c from below, the long range order $\overline{L}(B=0,T)$ is given by

$$\overline{L}(B=0,T) = (1-T/T_c)^{1/8} \quad (7)$$

while it is zero if T approaches T_c from above. It may be noted here that both the mean field and the Bethe approximation predict completely wrong behaviour i.e. they predict that $\overline{L}$ will go like $(1-T/T_c)^{1/2}$. These approximations do not also reproduce the logarithmic divergence of the specific heat. This difference with the approximate theory is rather important. One realizes that near T_c the nature of singularity is rather subtle. No matter what approximation one does, short of an exact solution, one never gets the log. This drives home a point that singularity is due to certain special circumstances that get killed by the approximations. It was several years later, that it could be identified as the effect of fluctuation, as reflected through fluctuation–response theorem. The universality of various phase transitions comes from the nature of fluctuations. It is the renormalization group approach developed by Kadanoff, Wilson and others, that gave the proper framework to handle fluctuations[21].

Unsolved problems

To appreciate how nontrivial Onsager's exact solution of the 2-dimensional Ising model was, let us remember that till today the exact solution of the 2-dimensional Ising model in nonzero external magnetic field has not been obtained. Further, the 3-dimensional Ising model with or without external magnetic field is also an unsolved problem. Similarly the 2-dimensional Ising model with any additional interaction (as, e.g., nearest and next-to nearest neighbour interactions) is also an unsolved problem. However special cases like the three-spin interaction on a triangular lattice can be solved[23].

5. Influence

Let us go back to the paper and discuss some of the developments where the 1944 paper played a direct role.

Duality

Duality between high temperature and low temperature for the Ising partition function was known before the Onsager Solution. However, Onsager used it to the fullest extent and introduced the idea of star-triangle (ST) transformation that could be used to solve the triangular lattice problem. The existence of duality is important because 'it always converts order into disorder and vice versa'[4]. Since the low temperature ordered phase is characterized by an order parameter, it seems possible to define an analogous quantity for the high temperature phase also. It is only in the eighties that such 'disorder' parameters played an important role in conformal invariance. The ST transformation turned out to be important in a different development when attempts were made in the 70's to solve more complex models, like vertex models. The existence of duality and ST guarantees the existence of commuting transfer matrices as expressed through the Yang–Baxter equation. The solvability of these models relies heavily on such commuting matrices[22].

Transfer matrix

The major problem was to diagonalize the transfer matrix. The largest eigenvalue determines the thermodynamic behaviour. Onsager could find all the eigenvalues for a finite lattice. This had implications far beyond the original idea of getting only the free energy.

Onsager's result showed that for a square (or any 2-dimensional) lattice, four different terms are needed if periodic boundary conditions are used. Why four was answered in a clear way by Kasteleyn in 1963 when he gave a general procedure for solving the Ising model using combinatorics ('dimers'). Kasteleyn showed that for a surface of genus g, 4^g 'terms' are needed. For open boundary conditions i.e. for strictly planar lattices, $g = 0$ while, for periodic boundary conditions, $g = 1$ as it has the geometry of a torus[23]. This also indicates that pushing the same approach to solve a 3-d lattice would require an infinite number of terms.

Correlation length and correlation

Since Onsager could find all the eigenvalues, he could also identify the correlation length. If $\lambda_1 > \lambda_2 > \lambda_3 \ldots$ are the eigenvalues, then the partition function can be written as

$$Q = \lambda_1^N \left[1 + \left(\frac{\lambda_2}{\lambda_1}\right)^N + \ldots\right], \quad (8)$$

where N is the number of rows in the direction of transfer. The phase transition occurs when the largest eigenvalue is degenerate. Close to the critical point, one can write $(\lambda_2/\lambda_1)^N \sim e^{-N/\xi}$ where $\xi = -(|\ln \lambda_2/\lambda_1|)^{-1}$. This identifies ξ as a special length scale to characterize the critical behaviour. The degeneracy of the eigenvalues as $T \to T_c$, in turn, implies a diverging length scale at the transition point. In this way, Onsager showed that $\xi \sim |T-T_c|^{-\nu}$ with $\nu = 1$. The renormalization group approach mentioned earlier is based on the idea of a diverging length scale so that the system is scale invariant at the critical point.

Later on, in 1949 Kaufman and Onsager showed that at $T = T_c$, spin correlation $\Gamma(r) \equiv < S(0)S(r) >$ decays as $r^{-1/2}$ for large distance[18]. In 1965, Patashinskii and Pokrovsky suggested the now famous scaling form $\Gamma(r,T) \approx r^{-1/4} D(r/\xi)$ where, apart from the power law prefactor, all distances are scaled by the correlation length at the temperature[24]. Explicit calculations could put the correlation function, in the proper asymptotic limit, in the scaling form but it was also noted that for $T < T_c$, $\Gamma \sim e^{-r/\xi}/r^2$. This anomalous behaviour in the low temperature phase is a very peculiar property of the Ising model in zero magnetic field. It is in the eighties that this point got clarified through the use of directed polymers[25].

Boundary effects

Onsager recognized the importance of boundary conditions. He realized that changing the couplings in one row to antiferromagnetic coupling ($J \to -J$) would increase the free energy in the ordered state because across this line spins cannot be aligned as in the bulk. This change, in other words, will force an interface in the system. The increase in free energy is then the energy of the interface or the surface tension. Since the interface vanishes at the critical point, the surface tension also vanishes at $T = T_c$. The exact solution showed that the surface tension vanishes as $(T_c - T)^{\mu}$ with $\mu = 1$. More general arguments later on proved that $\mu + \nu = 2$ for the 2-dimensional Ising model. For any other model the right hand side would involve the specific heat exponent also.

Finite size effects

Since the partition function was known for finite lattices, the specific heat could also be calculated. It was shown that, at $T = T_c$, the specific heat per spin, C, is finite. No phase transition can take place in a finite system. Furthermore, $C \sim \ln N$ as $N \to \infty$ (for a $N \to \infty$ lattice). It took many years to grasp the significance of this result until the idea of finite size scaling was introduced in 1970. Finite size scaling now plays an important role in the analysis of numerical data in a wide variety of problems.

A more dramatic result, though cannot be found explicity, in the Onsager paper, is the finite size behaviour of the free energy. Collecting various terms, from the calculation of the specific heat, it was found that for an $m \times n$ lattice

$$\lim_{m\to\infty} \frac{\ln Q_{mn}(T_c)}{mn} = \frac{2G}{\pi} + \frac{1}{2}\ln 2 + \frac{\pi}{12}\frac{1}{n} + \ldots, \quad (9)$$

where G is the Catalan constant. This is to be contrasted with eq. 8 for $T \neq T_c$ where there is an exponential approach to the large lattice (thermodynamic) limit. In 1986, Affleck argued using conformal invariance that in two dimensions, at criticality.

$$\lim_{m\to\infty} \frac{\ln Z}{mn} = \text{const} + \frac{c\pi c}{6n}$$

with c the central charge as the only quantity required to identify the critical behaviour ('universality class') in 2 dimensions[26]. A comparison of the two gives $c = 1/2$ for Ising model.

Experiment

The simplicity and wide applicability of the 2-dimensional Ising model and its solution still leave behind a queer sensation that, after all, the model is a bit artificial. So far as the exponents are concerned, they are universal and a wide variety of systems do have Ising exponents. It however remained a challenge to find an experimental system that could exhibit the behaviour one finds from a landmark of theoretical physics. The challenge was met when an alloy was found in 1974 which showed the behaviour of magnetization exactly like the Onsager result with only T_c as the adjustable parameter[27].

1. Based on talks given by two of us at Institute of Physics, Bhubaneswar on November 28, 1994.
2. *Phys. Today*, Feb. 1977, p. 77.
3. Two excellent review articles on Ising model are Newell, G. F. and Montroll, E. W., *Rev. Mod. Phys.*, 1953, **25**, 353–389; Brush, G., *Rev. Mod. Phys.*, 1967, **39**, 883–893. Whereas the first article gives an excellent technical discussion, the latter has a very good account of the historical developments; An excellent textbook on Ising model is by McCoy, B. M. and Wu, T. T., *Two-Dimensional Ising Model*, Harvard University Press, 1973.
4. Onsager, L., *Phys. Rev.*, 1944, **65**, 117–149.
5. Lenz, W., *Z. Phys.*, 1920, **21**, 613.
6. Ising, E., *Z. Phys.*, 1925, **31**, 253–258.
7. Lee, T. D. and Yang, C. N., *Phys. Rev.*, 1952, **87**, 410–419.
8. Heisenberg, W., *Z. Phys.*, 1928, **49**, 619–636.
9. Bragg, W. L. and Williams, E. J., *Proc. R. Soc.*, 1934, **A145**, 199; 1935, **A151**, 540.
10. Gorsky, W., *Z. Phys.*, 1928, **50**, 64.
11. Bethe, H. A., *Proc. R. Soc.*, 1935, **A150**, 552–575.
12. Peierls, R., *Proc. Camb. Phil. Soc.*, 1936, **32**, 471–476, 477–481.

13. Fisher, M. E. and Sherman, S., unpublished; Griffiths, R. B., *Phys. Rev.*, 1964, **A136**, 437–439.
14. Various papers in *J. Stat. Phys.*, **78**, Jan 1995, a special issue dedicated to Lars Onsager.
15. Kramers, H. A. and Wannier, G. H., *Phys. Rev.*, 1941, **60**, 252–262, 263–276.
16. Montroll, E., *J. Chem. Phys.*, 1941, **9**, 706; 1942, **10**, 61.
17. Kaufman, B., *Phys. Rev.*, 1949, **76**, 1232–1243.
18. Kaufman, B. and Onsager, L., *Phys. Rev.*, 1949, **76**, 1244–1252; for an overview see, Au Yang, H. and Perk, J. H. H., 1987, **A144**, 44–104.
19. Onsager, L., *Nuovo Cim.*, Suppl., 1949, **6**, 261.
20. Yang, C. N., *Phys. Rev.*, 1952, **85**, 809–816.
21. See, e.g., the articles in *Phase Transitions and Critical Phenomena* (eds Domb, C. and Green, M.), Academic Press, 1976, vol. 6.
22. Baxter, R. J., *Exactly Solvable Models*, Academic Press, 1982.
23. Kasteleyn, P., *J. Math Phys.*, 1963, **4**, 287–293.
24. Patashinskii, A. Z. and Pokrovsky, V. L., *Zh. Eksp. Teor. Fiz.*, 1966, **50**, 439–447 (English translation) *Sov. Phys. JETP*, 1966, **23**, 292–297.
25. Fisher, M. E., *J. Stat. Phys.*, 1984, **34**, 664.
26. Affleck, I., *Phys. Rev. Lett.*, 1986, **56**, 746–748.
27. Ikeda, H. and Hirakawa, K., *Solid Stat. Comm.*, 1974, **14**, 529.

MEETINGS/SYMPOSIA/SEMINARS

IUTAM-Symposium on Interaction Between Dynamics And Control In Advanced Mechanical Systems

Date: 21–26 April 1996
Place: Eindhoven, The Netherlands

Contact: Prof. D. H. van Campen
Chairman, Dept. of Mechanical Engineering
Eindhoven University of Technology
P.O. Box 513
NL-5600 MB Eindhoven, The Netherlands
Telefax: +31 40 447 355

IUTAM-Symposium on Innovative Computational Methods for Fracture and Damage

Date: 30 June–5 July 1996
Place: Dublin, Ireland

Contact: Dr. P. E. O'Donoghue
Chairman, Dept. of Civil Engineering
University College Dublin
Earlsfort Terrace
Dublin 2, Ireland
Telefax: +353 1475 4568

IUTAM-Symposium on Variable Density Low Speed Turbulent Flows

Date: 7–10 July, 1996
Place: Marseille, France

Contact: Professor L. Fulachier
Chairman, Institut de Mécanique Statistique de la Turbulence
12, Ave. Général Leclerc,
F–13003 Marseille, France
Telefax: +33 91 08 16 37
Phone: +33 91 50 54 39

Professor J. L. Lumley
Sibley School of Mech. and Aeros. Engin.
Upson and Grumman Halls
Cornell University
Ithaca, NY 14853, USA
Telefax: +1 607 225 12 22
Phone: +1 607 225 40 50

IUTAM-Symposium on Mechanics of Granular and Porous Materials

Date: 15–17 July 1996
Place: Cambridge, UK

Contact: Dr. N. A. Fleck
Chairman, University of Cambridge
Engineering Department
Trumpington Street
Cambridge CB2 1PZ, UK
Telefax: +44 1223 332 662
Phone: +44 1223 332 650

National Symposium on Polymers

Date: 26–27 February 1996
Place: Vallabh Vidyanagar

Topics include: Polymer synthesis and characterization; Polymer for specialized applications; Biodegradable polymers, environmental and ecological aspects of polymers; Polymeric dyes; Mathematical modelling in the field of polymers; Coordination polymers.

Contact: Prof. V. S. Patel
Coordinator, National Symposium on Polymers
Department of Chemistry, Sardar Patel University
Vallabh Vidyanagar 338 120
Phone: 02692–30411
Fax: 91–02692–46475

Introduction

David E. Bloom

Department of Economics, Columbia University, New York, NY 10027, USA
Harvard Institute for International Development, 1 Eliot Street, Cambridge, MA 01238, USA

The first case of HIV infection in India was detected in 1986 through a random blood test of a commercial sex worker in Madras. By late 1994, India's National AIDS Control Organization reported that nearly 51,000 HIV-positive people had been detected in the country, with 885 known cases of full-blown AIDS. Reported cases of HIV infection and AIDS, however, are only the tip of the iceberg and reflect little more than the limited amount of costly HIV testing that has been conducted. They are grossly misleading with respect to the extent of the epidemic, which needs to be estimated scientifically using statistical tools to infer the population level of infection from sample data. One such set of estimates recently prepared for India indicates that there are 900,000 to 2.1 million HIV-infected people in India, with 55,000 to 100,000 cumulative cases of AIDS. By the year 2000, the number of HIV infections is projected to grow to between 2.1 and 6.7 million people, with a cumulative total of 500,000 to 1.2 million cases of AIDS.

These statistics provide strong evidence that India is not immune to the rapid spread of HIV. It will not be spared because of its culture or through some ayurvedic or homeopathic cure. On the contrary, India is extremely vulnerable to the epidemic. Many factors contribute to its vulnerability: India's integration into the world economy; its changing social customs; its pockets of injecting drug use; its system of commercial blood supply; its medical practices; its large migrant population; its high level of illiteracy; and its age distribution, which has a large proportion of the population in the sexually active years.

AIDS is an unusual disease in modern times in that it is both incurable and invariably fatal. It is also unusual because carriers of HIV, the virus that causes AIDS, typically show no symptoms of the specific illnesses that define them as having AIDS for eight to ten years after they are infected. Thus, in its early stages, AIDS is an invisible disease. For this reason, in part, policy-makers are likely to be slow to react to the AIDS epidemic. It takes great wisdom, foresight and determination to address invisible problems, especially in the face of the many visible health problems, such as malaria, tuberculosis, hepatitis, cholera, typhoid and diarrhoea, which widely afflict many Asian populations. Responding to the AIDS epidemic is also complicated by the technical difficulty of establishing priorities for allocating resources to various health and other social problems, and by uncertainty about the appropriateness of particular policy and programmatic measures for HIV prevention and AIDS care.

Because of the disease's complexity in both biomedical and socioeconomic terms, no one discipline has a monopoly on the information and expertise needed to address the AIDS epidemic. Rather the response will have to draw upon skills and experiences from a wide range of fields, working cooperatively. This special issue of *Current Science* highlights this multidisciplinary complexity.

The nine articles that follow are divided into three groups. The first group provides essential background on HIV and AIDS, with John Dwyer discussing biomedical issues and James Chin describing the techniques for estimating and projecting HIV and AIDS cases from different sources of data. The second group describes the spread of HIV in Asia. Shiv Lal and B. B. Thakur provide an up-to-date review of the HIV and AIDS epidemic in India, while Rabin Sarda and George Peterson focus on the epidemic in the Western Pacific. Tim Brown and Werasit Sittitrai make the point that heterosexual sex is emerging as the dominant transmission category for HIV in Asia, and discuss salient features of the connection between risky heterosexual behavior and the epidemic.

The final group of articles addresses selected policy issues. Sherry Glied discusses the impact of HIV/AIDS from the perspective of the US experience, where nearly 300,000 people had died of AIDS by early 1995. Nandrani S. de Zoysa describes the Sri Lankan blood bank system and the measures it has adopted to safeguard the blood supply – almost entirely – from HIV contamination. Ronald Bayer discusses ethical considerations surrounding the issue of testing for HIV. Finally, Ajay Mahal and I discuss economic implications of HIV and AIDS and the potential contributions of economists to the design and implementation of policy and programmes in this area.

The authors hope that these articles will catalyse further research and discussion on the HIV and AIDS epidemic. India's natural and social science communities need to build momentum in their efforts to promote deeper understanding of the epidemic as a basis for the development of prevention and care strategies that are rooted in local realities. If the Indian epidemic continues along its present trajectory, the consequences could be disastrous. India must act now, so that it does not look back ten years from now with sorrow and regret upon 20 million HIV-infected Indians.

*I would like to thank the United Nations Development Programme's Regional Project on HIV/AIDS for supporting the publication and distribution of this special issue. I would also like to thank Preeti Sinha, Ajay Mahal and Uma Kurup for their advice and assistance throughout the project.

Biomedical aspects of HIV and AIDS

John M. Dwyer

Division of Medicine, Prince of Wales Hospital, Randwick, NSW 2031, Australia

IN 1981 reports of immunodeficiency associated with the development of pneumocystis carinii pneumonia in homosexual men on both the East and West coasts of the United States brought to the world's attention the beginning of what is now referred to as the AIDS epidemic[1]. Previously healthy men had, for no apparent reason, acquired an immune deficiency state that was so serious that they could no longer protect themselves from usually readily dismissed organisms such as pneumocystis carinii. Physicians in the Western world had not encountered such a syndrome before. Because the new disease initially affected homosexual men, concepts of a "gay plague" became popular and much effort was expended on looking for unique features of the homosexual life-style that might destroy immunological competence. Any retrospective review must realize that no knowledge of the epidemic of AIDS that had been sweeping through Africa for at least thirty years was available to help clinicians understand the disease they were now seeing in the West.

With the discovery that the human immunodeficiency virus (HIV) was responsible for destroying the immune system and producing AIDS[2] the pandemic nature of the problem became clear. At this writing, the World Health Organization's Global Programme on AIDS estimates that some eighteen million adults and one-and-a-half million children are infected with HIV and that between 5000 and 10,000 individuals are infected every 24 hours[3]. In 1994 more than three million adults became infected with the virus. The epidemic continues to spread at ever faster rates in many developing countries, with the epicentre of the epidemic moving inexorably from Africa to Asia. Already, India may have more individuals infected with HIV than any other country. This review will examine the biomedical aspects of HIV infection that are producing much individual human suffering and socioeconomic disruption in many countries.

Virology

Only in the last fifteen years have investigators realized that members of the retroviral family of viruses are capable of producing disease in humans[4]. Prior to that time the medical community had known that these viruses could produce disturbances of the immune and nervous systems of animals, with the immunodeficiency produced often being associated with some form of malignancy. A competent immune system is necessary to protect us from certain forms of cancer. We now know that one distinct branch of the retroviral family can produce leukaemia in humans while another branch of the family contains two viruses, HIV-1 and HIV-2, that can destroy the most important, indeed pivotal, cells within the immune system: CD4 lymphocytes. These cells are members of the T cell family of lymphocytes produced in the thymus gland. While the destruction of sufficient numbers of CD4 lymphocytes to produce immunodeficiency takes a number of years (there are approximately six hundred billion of such cells available to an adult), a threshold is reached when there are simply too few of these cells available to protect an individual from invading microorganism and the clinical syndrome AIDS develops.

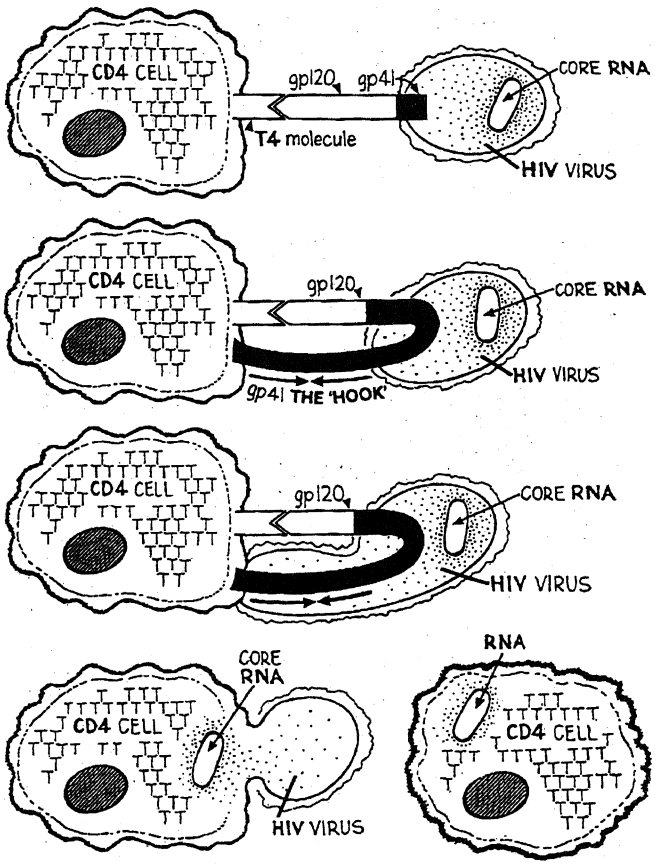


Figure 1. The binding and fusing of HIV to CD4 cells.

Retroviruses are composed of RNA and to reproduce must insert their genetic information into the DNA of the nucleus of the cell they invade. To accomplish this, they carry with them a unique enzyme, reverse transcriptase, that once activated inside a parasitized cell can translate RNA into DNA. It is this 'turning' from RNA into DNA that gives the family its name ('retro'). The virus is composed of an envelope from which protrude unique three-dimensional structures contained within a large molecule (molecular weight (MW) 120,000 Da) (Figure 1). This protein, gpl20, allows HIV to bind with exquisite specificity to a molecule present on the surface of CD4 thymus-derived lymphocytes. These cells are involved in initial recognition of 'foreignness' i.e. antigenic determinants which are recognized in the form of small amino acid sequences. CD4 lymphocytes do not react to self-determinants but do react to foreignness, initiating a cascade of immunological responses that ultimately lead to the elimination of non-self protein[5]. Currently, we recognize two distinct forms of CD4 lymphocyte. THE cells can stimulate cells referred to as CD8 lymphocytes (also derived from the thymus gland) to destroy directly or help produce an inflammatory response to viral, fungal and parasitic invaders. The second type of CD4 lymphocyte (TH2 cells) activate bone-marrow-derived lymphocytes, 'B cells', to produce antibodies that are essential for the elimination of extracellular bacteria[6].

Because CD4 lymphocytes play a pivotal role in initiating an effective immune response, the invasion and destruction of these cells by HIV-1 and HIV-2 cripples the entire immune apparatus.

As mouse lymphocytes into which the genes for the CD4 molecule have been inserted cannot be infected with HIV-1 or HIV-2, despite the resulting expression of the CD4 molecule on their surface, researchers have long argued that a second receptor is necessary for HIV invasion of lymphocytes. Recently scientists at the Institute Pasteur in Paris have suggested that the second molecule is CD26, a widely distributed protein present on many cells including CD4 lymphocytes[7].

A transmembrane protein in the envelope of HIV-1 and HIV-2 referred to as gp41, as it has a molecular weight of 41,000 Da, is capable of fusing the envelope of HIV with the membrane of a CD4 lymphocyte that has bound the viral receptor molecule gpl20. With the fusion of the two membranes, the core of HIV finds itself in the cytoplasm of the attacked CD4 lymphocyte. Activation of reverse transcriptase, insertion of proviral DNA and the subsequent replication of HIV within the T cell leads to both replication of the virus and the destruction of that particular T lymphocyte[8].

An addition to specific genes that code for the production of the viral envelope or the RNA sequences needed for replication and reverse transcriptase, HIV-1 and HIV-2 contain a series of genes that produce regulatory proteins. While some of the products of these genes may be directly toxic to CD4 lymphocytes, their major role is to activate and control the replication of HIV in an infected cell. These regulatory genes are potential targets for therapeutic intervention strategies. Thus, HIV has a unique capability to invade and destroy thymus-derived CD4 lymphocytes, thereby producing, over an extended period of time, such a gross immune deficiency that infected individuals can no longer limit the replication of HIV or respond effectively to other infectious agents that take the 'opportunity' presented to them to produce serious clinical disease.

Immunology

Infection with HIV produces, within six weeks, an effective immune response that limits the initial burst of viral replication. An effective response involves the production of neutralizing antibodies but, more importantly, the activation of CD8 lymphocytes that can destroy virus-infected cells and/or secrete an as yet uncharacterized lymphokine that inhibits viral replication. Lymphokines are hormone-like chemicals produced by activated lymphocytes to carry activation and suppressive signals from one cell involved in an immune response to another. Patients infected with HIV, though infectious, remain well for a number of years after infection occurs as a result of this immune response. No evidence exists to date to suggest that the immune response of anybody infected with HIV has successfully eliminated the virus or produced an immunological response that would make one confident that AIDS will never develop. While 67% of patients infected with HIV will have developed AIDS within ten years of infection, those that stay well longer are currently being investigated to determine if they make a more effective immune response to HIV than those who are ill at the end of ten years of infection.

Within a few weeks of infection the number of CD8 lymphocytes present in the circulation of people infected with HIV increases markedly. Most of the CD8 lymphocytes belong to that subclass of thymus-derived lymphocytes that are capable of attacking virus-infected cells. The total number of circulating CD4 lymphocytes usually undergoes an initial and significant fall around the time of seroconversion before stabilizing for many years. The replication of the virus in the first few weeks of infection is associated with permanent destruction of a significant number of CD4 lymphocytes.

HIV is cleared from the plasma of an infected individual at the time of seroconversion as antibodies remove the virus present and CD8 lymphocytes minimize further production. From then on, finding viral proteins in serum or indeed in peripheral blood lymphocytes becomes difficult. Thus, while it is possible throughout the course

of an active infection with the hepatitis B virus to find viral protein in the blood stream, we rely on the presence of antibodies to HIV to diagnose the presence of that virus. We now know that most of the infected CD4 lymphocytes pool rapidly within the lymphoid tissue of infected individuals. Lymphoid tissue is organized into nodes or glands that are scattered around the body at critical sites. The immune system prefers to fight its immunological battles in these highly organized centres. It is probable, but not established, that this pooling is associated with activation of signals that attract HIV-infected cells into lymphoid tissue. Certainly, a significant number of patients rapidly develop lymphadenopathy (swollen lymph glands) and recent studies strongly suggest that the destruction of most of CD4 lymphocytes by HIV-induced mechanisms occurs in lymphoid tissue rather than in the circulation[9].

Ultimately, the immune system's effectiveness in controlling HIV is compromised because of the ongoing destruction of more and more CD4 lymphocytes. Some four to five years after initial infection, some selective process, as yet uncharacterized, commonly seems to destroy THE CD4 lymphocytes more rapidly than those capable of stimulating an antibody response (TH2 cells). Thus, although the number of CD4 lymphocytes in the circulation may remain relatively stable, the number of cells capable of stimulating an effective cell-mediated immune (CD8) response continues to fall. This dramatic change in the immunobiology of the natural history of the disease is not associated with changes in an infected individual's clinical state.

Numerous reasons are known to be associated with the virus's capacity to evade an effective immune response and ultimately, in most cases studied, destroy effective immunity. HIV can destroy cells that are not directly

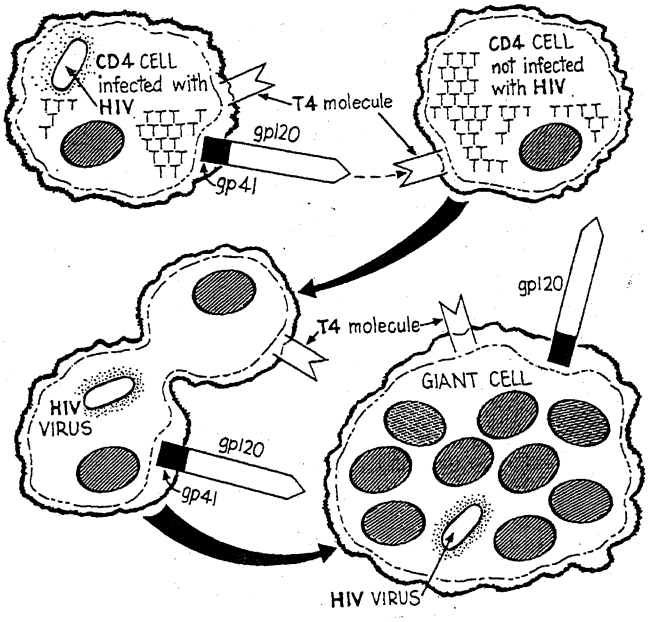


Figure 2. Two indirect methods used by HIV to kill CD4 cells.

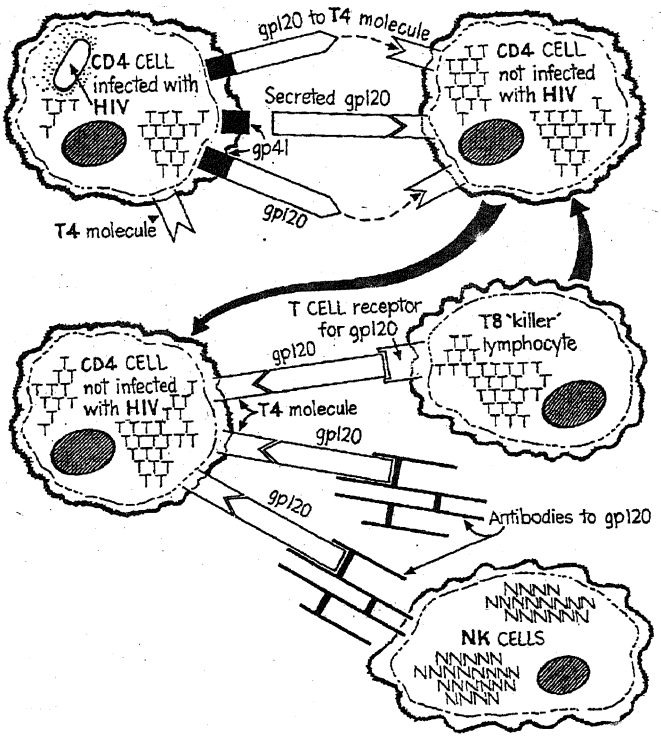


Figure 3. Normal CD4 (T4) cell destroyed by antibodies, NK or T8 cells because it bound secreted gp120.

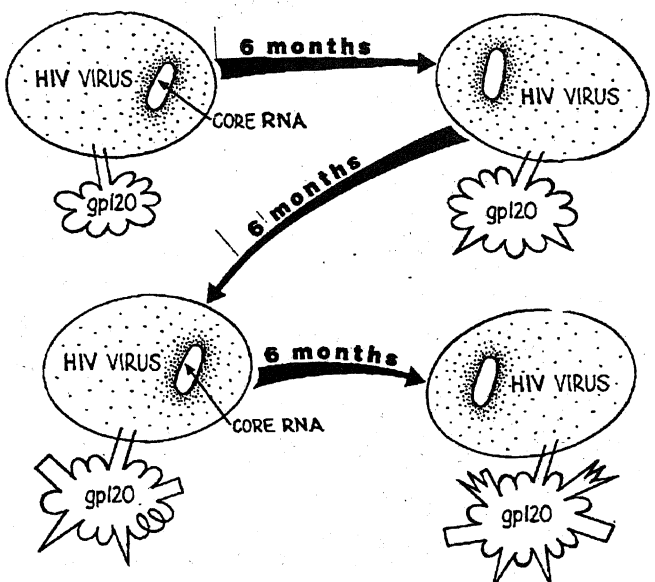


Figure 4. HIV changes the structure of its envelope (coat) to confuse attacking T8 cells.

infected with HIV at all (Figure 2). In cells infected with HIV in which the virus is actually replicating, envelope proteins (gp120 and gp41) are expressed on the surface of the cell, thus enabling an infected cell to bind the CD4 molecule of passing uninfected lymphocytes (Figure 3). As these envelope proteins contain the fusigenic moiety (gp41), giant cells (or syncitia) are formed. gp41 fuses the outer membrane of the T cell to the coat of the virus. At the

point of contact the contiguous membranes dissolve and in this way the core of the virus finds itself inside the CD4 lymphocyte. By the formation of these giant cells as infected cells trap uninfected cells it is possible for one infected cell to kill five hundred uninfected cells.

The secretion from infected cells of envelope proteins that will bind to CD4 molecules present on uninfected cells sets these unfortunate cells up as targets for cellular and humoral responses generated against envelope proteins. In addition to these strategies HIV-1 and HIV-2 are capable of generating considerable antigenic diversity among their envelope proteins. This constant antigenic 'drift' interferes with the perpetuation of an effective secondary immune response and requires the constant reactivation of primary immune responses to handle the variations presented to the immune system.

Clinical considerations

Infected CD4 lymphocytes, rather than free virions (elementary virus particles), are the important vector for HIV in human to human transmission. HIV-containing CD4 lymphocytes in sufficient concentration to make a biological fluid infectious are found in breast milk, blood and the sexual secretions of males and females. HIV can thus spread from an infected to an uninfected individual during vaginal or anal intercourse if infected blood from one individual enters the body of another and occasionally during breast feeding, particularly if a mother is in an early or very late stage of infection. At both extremes the replication of virus is most rapid and the concentration of virus in biological fluids reaches highest levels.

Within four to six weeks of infection, a seroconverting illness is noted by many, but these flu-like symptoms may not be particularly dramatic and are frequently not remembered by infected individuals. The symptomatology, which may occasionally feature significant pain associated with a neuropathy, is transient, and after it disappears many years may pass before an individual has any significant symptoms referable to HIV infection. Lymphadenopathy develops in 30–40% of infected individuals and this in itself is not a bad prognostic sign. As noted above, within ten years of infection with HIV, 67% of infected individuals have so many CD4 lymphocytes destroyed that they develop that ultimate complication of HIV infection: AIDS.

Prior to the development of AIDS, two HIV-related complications are common. The first is Kaposi's sarcoma, which is a tumour-like proliferation of the endothelial cells lining blood vessels[10]. The lesions often begin in the skin, although they are frequently first recognized in the mouth, particularly on the hard palate. The proliferation of these cells can occur within internal organs. Kaposi's sarcoma probably occurs when an as yet unidentified sexually transmitted cofactor stimulates endothelial cells to release cytokines which perpetuate the proliferation of these cells. Cytokines are like lymphokines, that are hormone-like substances released from cells, but in this case the chemicals referred to are released by cells lining blood vessels rather than lymphocytes. It is the limitation of Kaposi's sarcoma to homosexual men infected with HIV in the West that has led many to suggest that a sexually transmitted cofactor must be involved.

Kaposi's sarcoma can be cosmetically disfiguring but harmless, or can involve lymphatic tissue and produce gross swelling, particularly in the lower limbs and scrotum, and become a major clinical problem. Frequently, it displays a propensity for the nose and eyelids. Small lesions often respond to radiotherapy, but chemotherapy may be required if lesions are producing secondary phenomena of consequence.

The second complication is that some 10–20% of patients develop a presenile dementia prior to the development of AIDS. Initially, higher-centre functions are affected, such as concentration and judgment, and these defects can be measured using psychometric analysis. Antiretroviral therapy with AZT can minimize the symptoms of HIV-related dementia if introduced into the management of the condition soon after the onset of symptoms. This complication seems to be caused by certain strains of the virus that infect macrophages that transport the virus to the brain and then subsequently release cytokines that damage delicate nervous tissue. Macrophages are scavenger cells of the immune system and certain variations within the virus itself seem to make it easier for certain macrophages to swallow viral particles and transport them to the central nervous system.

Prior to the development of AIDS, as viral replication becomes more efficient, accelerated T cell destruction occurs, and patients may develop a period of weight loss characterized by profuse sweating at night and the development of a fungal infection in their mouth with *Candida albicans*.

CD4 lymphocytes are usually present in the blood in numbers greater than 700 µl. When the number falls to 200 µl or less, patients are at risk of developing opportunistic infections, which lead to the definition of AIDS in an individual case. Pneumonia produced by *Pneumocystis carinii* is the most common AIDS-defining illness. Toxoplasmosis, *Cryptococcus neoformans* infection (usually of the central nervous system), atypical and typical tuberculosis (the latter increasingly drug-resistant) and problems associated with cytomegalovirus are some of the most common serious infections noted. Once AIDS develops, patients usually survive no longer than two years even when the most sophisticated treatment is available.

Antiretroviral therapy

While numerous drugs to help manage HIV infection are in various stages of development, only one family of

drugs have so far been shown to be successful and relatively safe after extensive clinical trials. AZT was the first drug developed that could inhibit the replication of HIV effectively. As with all of the drugs that are currently available, AZT works by interfering with reverse transcriptase and by interfering with the successful transcription of viral DNA into viral RNA. Recent studies reveal that none of the currently available antiretroviral agents are effective long-term, as in most patients, viral mutations lead to drug resistance within 12–18 months. This appears to be true whether monotherapy or combination therapy is used. Strategies being studied now involve the rotation of various retroviral agents on a regular basis and the introduction of new drugs which attack viral replication at different stages of the virus's life cycle.

Vaccines

To date no vaccine has been developed that can prevent infection with HIV in humans or chimpanzees, the only animal besides humans that can be infected with HIV-1 or HIV-2. Although clinical trials in humans of preventative vaccines have started, no clinical data related to their efficacy is as yet available. Somewhat dishearteningly, evidence suggests that the vaccines used to prevent the infection of monkeys with a simian immunodeficiency virus (SIV) infection have been ineffective. Certain monkeys infected with that virus develop AIDS. Epidemiological investigations in the last twelve months have noted the emergence of more and more new strains of HIV in various parts of the world. There are now more than nine major variations on a theme which will make it extremely difficult to produce a vaccine that will be equally effective in different parts of the world. The difficulty in producing a cheap effective vaccine highlights the need to reemphasize preventative strategies. Because preventing infection is proving to be so difficult, many scientists have turned their attention to preventing disease. HIV infection does not produce significant clinical consequences in the first few years of infection. If the initially adequate immune response could be so strengthened that HIV replication is permanently minimized, then HIV infection may be no more serious than infection with the virus responsible for chicken pox. That virus is never eliminated from the body, but after an initial flourish rarely causes further problems. Therapeutic vaccines have been developed and are being tested in an attempt to introduce to a competent immune system viral envelope and core variations that the immune system may encounter at a later date during the course of HIV infection. While it is too early to judge the success of this strategy, on theoretical grounds it seems to be likely to produce more immediate benefits than any other therapeutic strategy currently being developed.

Summary

Two distinct members of the retrovirus family, HIV-1 and HIV-2, can selectively destroy CD4 lymphocytes in humans, and thus over a period of ten years can induce a degree of immunodeficiency that allows opportunistic infections and the development of malignancy to kill infected individuals. Envelope proteins developed by the virus provide numerous mechanisms for the destruction of immunologically competent cells. 80% of infections occur during vaginal intercourse, with blood-to-blood transmission being responsible for most remaining infections. At this time, anal intercourse is responsible for less than 10% of infections worldwide. No effective vaccine to prevent infection is likely to be developed within the next ten years. Prevention must remain the major goal of all public health officials, or the World Health Organization's prediction that 30–40 million people will be infected by the year 2000 may turn out to be a significant underestimate. Antiretroviral drugs and therapeutic vaccines offer the best hope in the immediate future for prolonging quality life in people infected with HIV.

1. Henry, M. and Michelis, M., *et al.*, An outbreak of community-acquired *Pneumocystis carinii* pneumonia, *New Engl. J. Med.*, 1981, **305**, 1431–1438.
2. Barre-Sinoussi, F. and Chermann. J., *et al.*, Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science*, 1983, **220**, 868–871.
3. *Global AIDS Newsletter*, 1994, **4.**
4. Pedersen, N. and Ho, E. *et al.*, Isolation of a T lymphotropic virus from domestic cats with an immunodeficiency syndrome, *Science*, 1987, **233**, 790–795.
5. Bottomly, K., A functional dichotomy in CD4+ T lymphocytes, *Immunol Today*, 1988, **9**, 268–273.
6. Nossal, G., B cell selection and tolerance, *Curr. Opin. Immunol.*, 1991, **3**, 193–198.
7. Gougeon, M. and Colizzi, V. *et al.*, New concepts in AIDS pathogenesis. *AIDS Res. Human Retrovir*, 1993, **3**, 287–289.
8. Seligmann, M. and Chess, L., *et al.*, AIDS – An immunologic reevaluation. *N. Engl. J. Med.*, 1985, **311**, 1286.
9. Pantaleo, G. and Graziosi, C. *et al.*, The immunopathogenesis of human immunodeficiency virus infection. *N. Engl. J. Med.*, 1993, **328**, 327–335.
10. Friedman-Kien, A., Disseminated Kaposi's sarcoma syndrome in young homosexual men. *J. Am. Acad. Dermatol.*, 1981, **5**, 468–471.

Estimation and projection of HIV infection and AIDS cases

James Chin

Division of Epidemiology, University of California, Berkeley, School of Public Health, Department of Biomedical and Environmental Health Sciences, 456 Kentucky Avenue, Berkeley CA 94707, USA

HOW many HIV infections and AIDS cases have occurred in the past, are currently present, and are likely to occur in the future? Reliable answers to these questions are needed to evaluate the spread of HIV and AIDS and to help plan public health, health care, and social welfare programmes vital for an appropriate response to this ever-expanding pandemic. However, since the early 1980s, when AIDS was first recognized, the pandemic's future trends and ultimate dimensions have been shrouded in uncertainty. This uncertainty persists because of the difficulties in measuring the prevalence, and especially the incidence, of AIDS cases and HIV infections in a given population with any substantial degree of precision. As a result, investigators have used many methods and models in attempts to understand the dynamics and interrelationships of the major determinants of HIV transmission and develop reliable estimates and projections of HIV and AIDS. The methods and models used need to be examined critically and understood before their outputs are accepted and used for public health programmes or policy decisions. This paper provides an overview of the uses and limitations of the methods available for HIV and AIDS estimates and projections. These methods are listed in Table 1.

Public health surveillance data[1–12]

Virtually all countries collect data on AIDS cases, primarily through public health surveillance systems that rely on passive notification of case reports. Reports of AIDS cases from industrial countries of Europe, North America and Oceania have been, up to 1993, based primarily on the 1987 Centers for Disease Control/World Health Organization (WHO) case definition[1]; those from Africa have generally been based on nationally adapted versions of the WHO clinical (Bangui) definition[11]; several countries in Latin America have used the Pan American Health Organization's Caracas definition[6]; and those from other countries involve a combination of these definitions. In addition to the problem posed by varying definitions of AIDS cases, the data are flawed by significant problems related to the wide variations found in the accuracy, completeness and timeliness of different national AIDS reporting systems. However, reported AIDS case data can provide the basis for estimating the actual number of cases that may have occurred. Researchers have prepared such estimates by adjusting reported AIDS data to correct for delays, inaccuracies and incompleteness. However, these adjustments require specific studies to determine the precise degree of correction needed for each limiting factor, and many estimators assume erroneously that these factors do not change with time. Hence, the reliability of AIDS estimates derived after such adjustments varies.

Detection of HIV infection became possible with the development of HIV antibody tests, which were first available in 1985. Most public health surveillance data on HIV infection have been obtained from a variety of uncoordinated sources, and much of the data were gathered from samples of selected population groups. Thus, data obtained from different areas or studies are usually not

Table 1. Methods for estimating and projecting of HIV infections and AIDS cases

Extrapolation from public health surveillance HIV/AIDS data
Delphi survey method
HIV and AIDS models
Type I models: statistical extrapolation from reported AIDS case data
Type II models: simple projection models
Type III models: complex deterministic/compartmental models
Combination of methods
Scenario/modelling approach

comparable. Current serological tests for detecting HIV antibodies are reliable for measuring HIV prevalence in populations: collectively they have shown that HIV infection is not randomly distributed in a given population. Relatively high prevalence rates (from a few per cent to more than 50) can be found among homosexual or heterosexual individuals who have had multiple sexual partners and those people who share drug injection equipment. The prevalence of HIV infection in most general populations of industrial countries is low and may range from almost zero to one per several thousand.

Researchers have estimated HIV prevalence most often by using the results of serological surveys and extrapolating the data to specific populations. Major problems with this method are the limited number of seroprevalence studies that are representative of specific populations or subgroups, and the wide variability in estimates of the size of important subgroups, such as female prostitutes, injecting drug users, and patients seen in sexually transmitted disease clinics. Nevertheless, investigators have derived reasonable working estimates of the prevalence, general distribution, and trends of HIV infection for many countries by analysis of all available HIV serosurvey data.

In the mid- to late-1980s, with only limited HIV data available, investigators tended to overestimate HIV prevalence for many areas and populations. In the United States for example, during the period 1985–1992 official estimates of cumulative HIV infection remained consistently in the range 1.0–1.5 million. As of early 1995, the best estimate of cumulative HIV infections was still between 1.0 and 1.5 million, which clearly indicates that all previous estimates were too high. In the United Kingdom in the mid-1980s HIV prevalence was estimated at 80,000. Subsequent reviews and analyses of the accumulating HIV data by expert committees resulted in a consistent downward revision of this estimate to a maximum of about 30,000. In 1987, WHO's estimate of global HIV infection ranged from 5 to 10 million. With the collection and analysis of more recent HIV data (primarily from sub-Saharan Africa) WHO's revised 1990 estimate of global HIV infection was reduced to the lower level of the 1987 estimate, i.e. about 5–6 million. Since the late 1980s, WHO's global and regional HIV estimates have been intentionally conservative.

The Delphi survey method[13–15]

The Delphi survey method was developed in an attempt to improve the reliability of the judgments needed in relatively uncertain situations and to provide a means of quantifying such judgements. This method was originally systematized in the 1960s for use in developing management policies. It was first applied to the health field in the 1970s. Essentially, the Delphi method obtains educated guesses from selected experts in a reiterative fashion, and then uses the average and range of the Delphi responses as projections. Major advantages of the Delphi method are speed and low cost. However, it is difficult to select truly knowledgeable experts to estimate the number of HIV infections and AIDS cases. Furthermore, estimates and projections made by the Delphi method may have extremely wide ranges.

In 1988 WHO used a Delphi survey to obtain projections of HIV prevalence. Delphi survey respondents projected that there would be 15–20 million cumulative adult HIV infections worldwide by the year 2000. Yet shortly after that survey, HIV serological data for sub-Saharan Africa and southeast Asia showed marked increases in HIV seroprevalence. These increases, estimated at about 5 million new infections within two to three years after the survey, suggested that the 1988 Delphi projections were too low. In 1991, based on observed HIV trends, WHO projected that a cumulative total of 30 to 40 million HIV infections would occur in men, women, and children by the year 2000.

In mid-1992 a Harvard group, AIDS in the World (AIW)[15], also carried out a Delphi survey to project cumulative global HIV infections by the year 2000. While AIW's low Delphi total was only slightly higher than WHO's conservative projection in 1991, its high Delphi projection, more than 100 million cumulative HIV infections by the year 2000, was three to four times higher than their low Delphi projection. With the exception of AIW's high Delphi projection (for which no comparable WHO projection was made) the independent estimates and projections of HIV infections and AIDS cases were reasonably close to WHO's projections. AIW did not select either the low or high Delphi projection as their best estimate for the year 2000, but most references to its Delphi projections focus on their high Delphi projection.

How plausible are AIW's high Delphi projections? If cumulative infections in North America (primarily the United States) were estimated to be about 1.2 million as of 1993, then almost 1 million new HIV infections would have to occur each year in North America to reach AIW's high Delphi projection of more than 8 million. Similarly, in sub-Saharan Africa, more than 3 million new HIV infections would have to take place annually to reach the projected total of more than 33 million by the year 2000. For other regions, the AIW's projections require new HIV infections of about 1 million per year in Latin America, close to 1 million per year in the Caribbean, and almost 6 million per year in southeast Asia to reach 45 million by the year 2000 (Table 2). None of these annual incidences are plausible based on observed HIV data and trends, which suggest that with the exception of southeast Asia, which is still at an early stage

Table 2. AIDS in the world: HIV estimates and Delphi projections (millions)

Region	Cumulative estimate Jan 1992	Cumulative estimate Jan 1993	High Delphi projection by the year 2000	Annual incidence needed from 1993 to 2000
North America	1.167	1.261	8.150	0.984
Western Europe	0.718	0.834	2.331	0.214
Sub-Saharan Africa	7.803	10.985	33.609	3.232
Southeast Asia	0.675	3.567	45.059	5.927
Latin America	0.995	1.050	8.554	1.072
Caribbean	0.310	0.461	6.962	0.929
Total	11.799	18.311	104.665	12.336

in its HIV epidemic, the highest annual HIV incidence in most other regions may have already occurred.

HIV and AIDS models

Researchers have developed many types of HIV and AIDS models in attempts to understand the dynamics of HIV spread and to develop reliable projections of the epidemic's eventual extent. The major components of all models include the basic assumptions and input parameters used, the value or range of values for the input parameters, and the mathematical interrelationships between the parameters used in the model. The results produced by any model should not be used to formulate programmes or make policy decisions unless those making such decisions fully understand and agree with the basic assumptions, the values used for the input parameters and the mathematical interrelations between the input variables.

HIV and AIDS models can be broadly classified into three types that range from the simplest to the most complex. The first two types are empirical, extrapolation models, while models of the last type have been described as explanatory or deterministic. In general, the simpler extrapolation models are designed for short-term projections and the more complex models for evaluating determinants of the HIV and AIDS pandemic and for hypothesis testing.

Type I models[16–18]

Type I models use reported AIDS case data to make short-term (two to three years) projections of AIDS cases. Such short-term projections have been made by statistical extrapolation and regression techniques applied to the observed temporal curve of the reported AIDS cases. These models assume that after adjusting for reporting delays (and in some models for incomplete reporting), the trends of reported cases during the next few years will remain essentially similar to those observed in the recent past. A profusion of such models has been described, but a major problem with these models is that many mathematical curves can fit the data equally well. Not only is there no way to choose the best among the various curves, but the fitted curves may lead to widely divergent projections, particularly if they are used to make projections for periods longer than two to three years. Type I models should be used only to project AIDS cases two to three years ahead, and only for populations for which AIDS case reporting is relatively timely, reliable and complete.

Type II models[19,20]

Type II models use the data on estimated HIV infections and progression rates from infection to AIDS to calculate the number of past AIDS cases and to provide short-term (three to five years) projections of AIDS cases. A variation of this approach is the back calculation method, which uses AIDS case reports in conjunction with annual progression rates from infection to AIDS to estimate the number of annual HIV infections that have occurred. This method requires relatively accurate, complete and timely reporting of AIDS cases. However, because few people develop AIDS within two to three years of infection, this method cannot be used to estimate HIV infections that occurred during the prior two to three years even in areas where AIDS case reporting is considered reliable.

A simplified variation of this method uses an estimated ratio of HIV infections to AIDS cases to calculate total HIV infections. Like back calculation, this ratio method requires reliable estimates of AIDS cases, which are not available in most developing countries. In addition, users of the ratio method must realize that the ratio of HIV infection to AIDS cases changes rapidly with time. The ratio falls from many thousands to one during the first few years of extensive HIV spread in a population to less than ten to one after the first decade. This decline occurs whether HIV incidence is increasing or decreasing because a substantial proportion of HIV-infected individuals progress to AIDS as long as 10 to 20 years after contracting HIV infection.

Type III models[21–26]

Type III models are more complex and incorporate biological and behavioural variables that describe the transmission and natural history of HIV infection so as

to simulate the entire disease process. Many mathematically sophisticated type III models have also been developed to project the future course of HIV and AIDS epidemics in different areas and populations. These complex models usually use differential equations to describe biological and behavioural variables known to be important in the transmission and progression of HIV disease in their simulation of the process from infection to progression to AIDS and death.

These deterministic or simulation models include: the mathematical model(s) of Anderson *et al.*[21]; the demographic model developed by Bulatao of the World Bank[25]; a Monte Carlo type model ('Simulaids') developed by Auvert *et al.*[22]; (1990); and an extremely detailed model – the interagency working group model – developed principally by Stanley and Seitz[26]. These models require, in varying degrees, extensive data sets on virtually all the demographic, biological and behavioural variables considered to be important in the epidemiology and natural history of HIV infection. The major problem with these models, aside from their complexity, is that most of the precise, detailed data sets they require to make projections are not available, even in those countries with the best data collection systems.

Most complex type III models have used hypothetical populations to obtain projections. In an exercise for which the population characteristics and a defined set of variables and input values were provided, a group of models produced widely varied, long-term demographic projections for HIV and AIDS. Projections for the percentage of the adult population infected with HIV in the year 2010 varied from a high of almost 40% in one model to a low of less than 3% in another. The greatest value of complex mathematical models may well be in their use to test hypotheses rather than for estimation and projection of HIV and AIDS prevalence.

A scenario/modelling approach

Given that predictions by complex models appear to be of doubtful validity, and because no criteria exist to select a good or best HIV and AIDS model (especially for projecting future HIV incidence and prevalence), a simpler scenario/modelling approach for estimation and projection of HIV infections and AIDS cases was developed by the author when he was at WHO during the late 1980s. The scenario/modelling approach can be used in individual countries and for selected populations within countries to provide working estimates and short-term projections of HIV-related morbidity and mortality for purposes of policy development and public health planning.

A scenario is defined as an outline of any series of events, real or imagined. Plausible HIV scenarios can be constructed from available HIV data and observed trends. Because future AIDS cases and deaths depend on the number and pattern of HIV infection and progression rates from infection to severe disease and subsequent death, AIDS cases and deaths can be calculated for different HIV prevalences and patterns (HIV scenarios) based on simple assumptions about the occurrence of these events.

The general methods used in the scenario/modelling approach to develop working estimates and projections of HIV and AIDS are as follows:

* The most recent patterns, prevalence and trends of HIV infection in specific population groups are estimated from an epidemiological analysis of all available HIV data. The HIV database compiled by the Center for International Research, US Bureau of the Census, is a valuable data source. Unpublished national HIV data, when available, should also be used to supplement these data.
* Based on these data and other epidemiological observations, different HIV scenarios can be estimated and projected to the year 2000 or beyond for major geographical and epidemiological regions. Optimistic or best-case scenarios should be constructed to derive conservative estimates and projections for public health planning.
* EPIMODEL, a relatively simple type II AIDS model developed by the author when he was with the WHO's Global Programme on AIDS can be used to estimate and project AIDS cases and deaths based on the HIV scenarios.

The reliability of HIV and AIDS projections obtained by this scenario/modelling approach depends primarily on two factors: the accuracy of the estimated number of HIV-infected people, and the proportion of HIV-infected people who ultimately develop AIDS.

The estimates and projections obtained from this scenario/modelling approach need to be modified as additional data on HIV prevalence and progression indicate that changes are warranted. However, health and social welfare systems should use conservative estimates and projections derived by this method to plan for and develop the resources needed to respond to the minimum number of HIV-related disease and deaths that can be expected during the remainder of this decade.

Summary and conclusions

The answer to the opening question 'How many HIV infections and AIDS cases have occurred, are currently present, and will occur in the future?' is that we really do not know, but we can estimate and project these numbers with varying degrees of uncertainty. Our ability to estimate past and current numbers is improving as public health surveillance systems collect more data.

Although the accuracy of such estimates varies greatly depending upon the quantity and quality of the surveillance data, plausible minimum, maximum and mid-range working estimates can be developed for use by public health and health care programmes.

Projection of future patterns and prevalence of HIV and AIDS is more difficult, but such projections or scenarios are essential to begin planning programmes to respond to the increasing number of premature deaths among young and middle-aged adults because of AIDS. A conservative HIV and AIDS scenario can be constructed from data on current global status and trends of regional epidemics. According to this scenario a cumulative total of about 40 million men, women and children would have been infected with HIV by the year 2000, with more than 90% of this total occurring in developing countries. By understanding the uses and limitations of all available methods and models that have been used for forecasting HIV and AIDS epidemics, such scenarios can assist policy-makers and social welfare and health care planners to prepare for the projected numbers of AIDS deaths and AIDS-related orphans.

1. Centers for Disease Control, Revision of the CDC surveillance definition for acquired immunodeficiency syndrome, *Morbid. Mortal. Weekly Rep.*, 1987, **36** (supplement 1S).
2. Chin, J., Public health surveillance of AIDS and HIV infections, *Bull. WHO*, 1990, **68**, 529–36.
3. Chin, J. and Mann, J. M., The global patterns and prevalence of AIDS and HIV infection, *AIDS 1988*, 1989, **2**, S247–S252.
4. Chin, J. and Mann, J. M., Global surveillance and forecasting of AIDS, *Bull. WHO*, 1989, **67**, 1–7.
5. Department of Health and the Welsh Office, Short-Term Prediction of HIV Infection and AIDS in England and Wales, Report of a Working Group (ISBN 011 321191 0), London.
6. Pan American Health Organization, Working Group on AIDS Case Definition, *Epidemiol. Bull. Pan Am. Health Organiz.* 1990, **10**, 9–11.
7. Public Health Laboratory Service, Acquired Immune Deficiency Syndrome in England and Wales to End 1993: Projections Using Data to End September 1989, Communicable Disease Report, (ISSN 0959-2547) pp. 1–12.
8. Public Health Laboratory Service, The Incidence and Prevalence of AIDS and Other Severe HIV Disease in England and Wales for 1992–1997: Projections Using Data to the End of June 1992, Communicable Disease Report 3 (ISSN 0264-1607) (Supplement 1): S1–S17.
9. Sato, P. A., Chin, J. and Mann, J. M., Review of AIDS and HIV infection: Global epidemiology and statistics, *AIDS 1989*, 1990, **3**, S301–S307.
10. United States Bureau of the Census, Center for International Research, AIDS/HIV Surveillance Database, Washington, DC, 1992.
11. World Health Organization, Report of a Workshop on AIDS in Central Africa. Bangui, Central African Republic, October 22–25, WHO/CDS/AIDS/85.1, 1985.
12. WHO, Global Programme on AIDS, Current and Future Dimensions of the HIV/AIDS Pandemic: A Capsule Summary, Geneva, 1992.
13. Chin, J., Sato, P. A. and Mann, J. M., Projections of HIV infections and AIDS cases to the year 2000, *Bull. World Health Organiz.*, 1990, **68**, 1–11.
14. Levine, A., A model for health projections using knowledgeable informants, *World Health Statist. Quart.*, 1984, **37**, 306–317.
15. Mann, J. M., Tarantola, D. J. M. and Netter, T.W., (eds), *AIDS in the World*, Harvard University Press for the Global AIDS Policy Coalition, Cambridge, MA, 1992.
16. Dondero, T. J., St Louis, M., Anderson, J., Peterson, L. and Pappaioanou, M., Evaluation of the estimated number of HIV infections using a spreadsheet model and empirical data, Fifth International Conference on AIDS, Montreal, Canada, June, 1989 (Abstract M.A.O.4).
17. Healy, M. R. J. and Tillett, H. E., Short-term extrapolation of the AIDS epidemic, *J. R, Statist. Soc.*, 1988, **151**, 50–65.
18. Morgan, W. M. and Curran, J. W., Acquired immunodeficiency syndrome: Current and future trends, *Public Health Rep.*, 1986, **101**, 459–465.
19. Brookmeyer, R. and Gail, M. H., A method for obtaining short-term projections and lower bounds on the size of the AIDS epidemic. *J. Am. Statis. Assoc.*, 1988, **83**, 301–308.
20. Chin, J. and Lwanga, S. K., Estimation and projection of adult AIDS cases: A simple epidemiological model, *Bull. WHO*, 1991, **69**, 399–406.
21. Anderson, R. M., May, R. M., Boily, M. C., Garnett, G. P. and Rowley, J. T., The spread of HIV-1 in Africa: Sexual contact patterns and the predicted demographic impact of AIDS, *Nature*, 1991, **352**, 581–589.
22. Auvert, B., Moore, M., Bertrand, W. E., Beauchet, A., Aegerter, P., Lusamba, D., Diong, K. T., and Klinowski. J., Dynamics of HIV infection and AIDS in central African cities, *Int. J. Epidemiol.*, 1990, **19**, 417–428.
23. Way, P.O. and Stanecki, K., The demographic impact of an AIDS epidemic on an African country: Application of the IWGAIDS model, Center for International Research Staff Paper No. 58, US Bureau of the Census, Washington, DC, 1991.
24. United Nations/World Health Organization, The AIDS epidemic and its demographic consequences, Proceedings of the United Nations/World Health Organization Workshop on Modelling the Demographic Impact of the AIDS Epidemic in Pattern II countries: Progress to Date and Policies for the Future, New York, 15–16, December 1989, United Nations Publication Sales No. E.91.XIII.5ST/ESA/SER.A/119, New York, 1991.
25. Bulatao, R. A., Projecting the demographic impact of the HIV epidemic using standard parameters, (Author/JC's note: Randy Bulatao participated in the above workshop and had his paper published in the above proceedings: pp. 90–104), 1991.
26. Stanley, E. A., Seitz, S. T., Way, P. O., Johnson, P.D. and Curry, T.F., The interagency working group (IWG) model for the heterosexual spread of HIV and the demographic impact of the AIDS epidemic, (Similar note as above: pp 119–136), 1991.

The problem of HIV and AIDS in India

Shiv Lal and B. B. Thakur

National AIDS Control Organization, Ministry of Health and Family Welfare, IRCS Building, 1 Red Cross Road, New Delhi 110 001, India

ALTHOUGH the AIDS pandemic is in its early stages in Asia, the virus is spreading rapidly among high-risk groups and is making fast inroads into the general population. Estimates suggest that more than one million people are infected with HIV in India[1].

India is a large country in terms of area and population, so the epidemic's spread and impact have not been homogenous . There are many regional and population group variations, which probably reflect separate epidemics with their own starting points and probably involve different types and frequencies of risky behaviours and practices, such as having multiple sex partners or sharing drug injection equipment.

The situation in India

HIV was first detected in India in 1986 when pilot HIV surveillance activities were carried out among selected populations and in various locations to ascertain if the virus was present in the country. Within six or seven years, it became evident that HIV infection was present in almost all parts of the country, but was worst in the states of Maharashtra, Tamil Nadu and northeastern States of Manipur, Mizoram and Nagaland. In the latter three states, which border Myanmar, the main mode of HIV transmission is through needle sharing by injecting drug users[2], whereas in the rest of the country, most infections are transmitted through heterosexuals[3]. The number of AIDS cases reported from all parts of the country is increasing.

AIDS cases

In 1986, Bombay reported India's first AIDS case. Since then, as of 31 January 1995, a total of 1032 AIDS cases have been reported to the National AIDS Control Organization from the states and union territories as shown in Table 1 (ref. 4).

As is true elsewhere, the number of reported cases in India represents a small fraction of actual AIDS morbidity. This gap between actual and reported cases is caused by both underreporting and misdiagnosis; as the medical community's awareness and training improve, the gap is likely to shrink. According to estimates based on HIV prevalence, the actual number of AIDS cases in India is probably between 10,000 and 20,000 (ref. 5). However, the number of AIDS cases in the country cannot be taken as indicative of the current HIV and AIDS problem because the number of current AIDS cases reflects the HIV situation that prevailed five to ten years ago.

Table 1. National AIDS Control Programme, India, AIDS cases in India (Reported to NACO, as of 31 January 1995)

S. No.	State/Union Territory	AIDS cases
1.	Andhra Pradesh	1
2.	Assam	2
3.	Delhi	68
4.	Goa	12
5.	Gujarat	18
6.	Haryana	1
7.	Himachal Pradesh	9
8.	Jammu & Kashmir	2
9.	Kerala	76
10.	Madhya Pradesh	21
11.	Maharashtra	288
12.	Manipur	68
13.	Pondicherry	6
14.	Punjab/Chandigarh	47
15.	Rajasthan	1
16.	Tamil Nadu	345
17.	Uttar Pradesh	8
18.	West Bengal	30
19.	Karnataka	26
20.	Dadar Nagar & Haveli	1
21.	Orissa	2
	Total	**1032**

HIV infection

From October 1985 until the end of January 1995, about 2,443,141 individuals had been tested for HIV through the nationwide surveillance, of whom 17,283 tested positive for HIV (ref. 4). HIV prevalence increased from 2.5

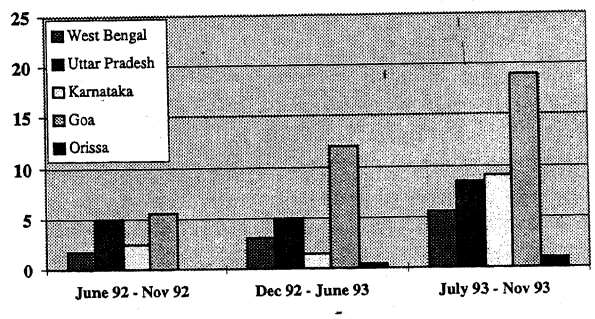


Figure 1. Prevalence of HIV infection in selected states (June 92–November 1993).

per 1000 population in 1986 to 11.2 per 1000 by 1992. The course of the epidemic has been quite varied in nature, as is reflected clearly by Figure 1. Most of this testing has been done on population groups perceived to be at high risk for HIV infection; this may influence the results. The data available to date indicate not only that the infection is present in all regions but also that it has spread well beyond the high-risk groups to the general population.

The HIV prevalence rate varies widely in different areas and among different population groups. The major concentration of infection remains in the city of Bombay, where HIV prevalence among prostitutes has risen significantly, from 10% in 1986 to 51% in 1993. The other known centres of HIV infection are Madras, Madurai, Pune, Vellore and the surrounding areas. In Vellore, the HIV rate among prostitutes increased from 0.5% in 1986 to 34.5% in 1990. The latest sentinel surveillance data (1993) from Madurai and Madras showed that the prevalence of HIV infection among attendees at clinics for sexually transmitted diseases is 6.5% and 2.7%, respectively. Other studies have shown that HIV prevalence among injecting drug users increased from some 45% in the first half of 1992 to 73% in the first half of 1993.

The increasing trends of HIV infection among the high-risk groups have been followed closely by the rising trends among the general population. The sentinel surveys done among pregnant women in Bombay, Salem and Manipur have found the prevalence of HIV to be 2.5%, 1.1% and 0.8%, respectively (Figure 2). The prevalence among pregnant women in Tirupati rose from 0.5% in 1990 to 2.0% in 1993. Such trends demonstrate the spread of HIV infection in the general population.

The trends in HIV prevalence rates among high-risk groups in India seem to be comparable to trends previously observed in some African nations among prostitutes. This similarity suggests that the epidemic could spread to the general population to alarming levels, as happened in these countries, where the spread of HIV among the general population seems to have been preceded by a sharp increase in HIV prevalence among high-risk groups.

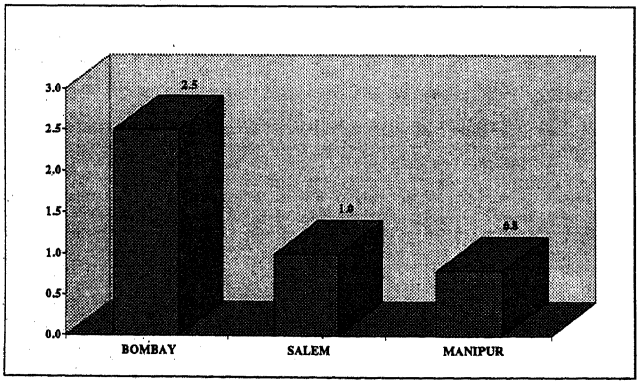


Figure 2. HIV positivity among pregnant women (1st round-sentinel surveillance.

Clearly, HIV infection is not restricted to such marginalized groups as prostitutes and drug addicts, but is spreading to more areas and to the general population in those areas through heterosexual transmission. Nevertheless, not all regions are being affected to the same extent. The apparent geographical clustering at this stage of the epidemic may be caused by differences in the proportion of the population engaging in high-risk behaviour or by unevenness in the surveillance system. Whatever the local pattern, however, the surveillance data taken as a whole demand that there be implementation of a comprehensive AIDS control programme as soon as possible.

Other issues

In India the prevalence of tuberculosis infection is 40% as evinced by tuberculosis testing. Of reported AIDS cases, 56% had tuberculosis of one or more organs. Thus, given this high prevalence of tuberculosis among those with AIDS, India is going to face the added problem of an increase in cases of tuberculosis even among those not suffering from AIDS.

Conclusion

Poverty and underdevelopment are key factors that contribute to the AIDS pandemic: financial difficulties and economic problems force men to leave their families to find work, encourage drug use and make prostitution a survival strategy[6]. If we take these factors into account, India has great reason to be alarmed. India recognizes this threat clearly. Accordingly, India has launched a comprehensive, nationwide AIDS control programme since 1992 with financial assistance from the World Bank and technical input from the World Health Organization. This programme consists of different components, namely: surveillance for HIV/AIDS; control of sexually transmitted diseases; condom distribution programmes; and information, education and communication designed to generate awareness of the disease and encourage change in risky behaviours. Further, the programme aims at providing comprehensive, nonstigmatizing care to all those who are already suffering from AIDS.

1. National AIDS Control Programme, India, Country Scenario – An Update, National AIDS Control Organization, Ministry of Health & Family Welfare, Government of India, 1993, pp. 7–9.
2. Sarkar, S., Rapid spread of HIV among injecting drug users in North Eastern states of India, *Bull. Narc.* 1993, **45**, 91–104.
3. Lal, S., Current Status of AIDS & HIV in India, *J. Indian Med. Assoc.*, 1994, **92**, 12–13.
4. Surveillance for HIV Infection in India, Monthly Report, National AIDS Control Organization, Ministry of Health & Family Welfare, Government of India, Feb. 1995
5. Lal, S., Khodekovich, L. and Sengupta, D., AIDS case surveillance, *J. Med. Assoc.*, 1994, **92**, 55–56.
6. Richard, H., A global agenda for AIDS, *Lancet*, 1993, **341**, 234.

The HIV and AIDS epidemic in the Western Pacific

Rabin M. Sarda and George Petersen*

WHO, United Nations Avenue, PO Box 2932, 1099 Manila, Philippines
*WHO, House 120, Street 228, Sankat Chadomuk, Khan Mon District, Phnom Penh. Cambodia

THE Western Pacific Region of the World Health Organization (WHO) consists of 35 countries and areas, with a total estimated population in 1992 of 1535 million[1]. Seven countries account for 96% of this total. The region includes industrial countries, newly industrializing countries or areas, developing countries or areas, and some of the world's least developed countries. The regional population is expected to grow to 1700 million by the year 2000. The average national growth rate was 2% per year during 1985–90.

Global trends

A cumulative total 1,025,073 AIDS cases had been reported[2] worldwide up to the end of 1994. WHO estimated that the actual figure was closer to 4.5 million cases. Of the reported cases, the vast majority have been from the Americas and Africa, with Asia and the Pacific accounting for less than 2% of the total.

WHO estimates indicate that up to the end of December 1994, more than 18 million people had been infected by HIV[2], with the majority of the new infections during that year occurring in Africa and Asia. By the year 2000, WHO predicts 30–40 million HIV infections worldwide, with 90% of them expected to occur in developing countries[3].

Current situation in the Western Pacific Region

Available evidence suggests that the HIV epidemic in the Western Pacific Region is at a relatively early state compared to other parts of the world[4]. Few countries in the region had reported significant numbers of AIDS cases or HIV diagnoses prior to 1987. Since then, the epidemic has gradually spread throughout the region. Only eight of the thirty-five countries and areas (predominantly the smaller Pacific island nations) have yet to report any case of AIDS or HIV infection.

As Table 1 shows, as of 1 February 1995, 7349 cases of AIDS had been reported[4] from countries in the region. Eighty-six per cent of AIDS cases have been reported from three countries: Australia, Japan and New Zealand, but this proportion is decreasing (Figure 1). Of the reported AIDS cases, the vast majority were men (94.1%) and less than 1% of the reported AIDS cases were children under 15 years of age[5].

A cumulative total of 39,386 HIV infections had been reported[6] in the Western Pacific Region as of 1 February 1995 (Table 1). Reported rates of HIV infection vary from 101 per 100,000 population in Australia and 64 per 100,000 in French Polynesia, to fewer than 9 per 100,000 reported in Hong Kong and 0.1 per 100,000 in China. Eighty-six per cent of infections in the region have occurred in men and 8% in women. The proportion of HIV infections among children remains small, less than 1% of the total.

New HIV infections in Australia have remained relatively constant in the past three years. In Japan, they seem to have peaked in 1992. In Malaysia, however, they have been increasing (Figure 2).

Reported cases of HIV infection or AIDS do not show the real prevalence of these conditions in any given population. For example, many who are infected with HIV

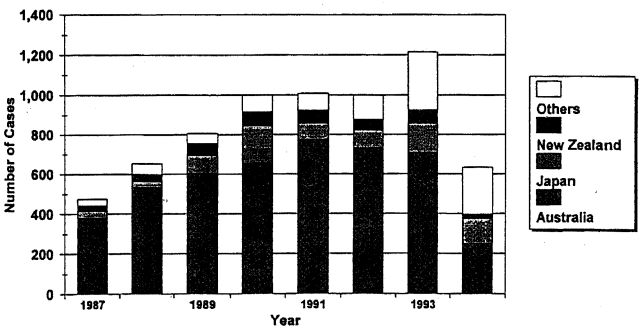


Figure 1. Annual reported incidence of AIDS, Western Pacific Region.

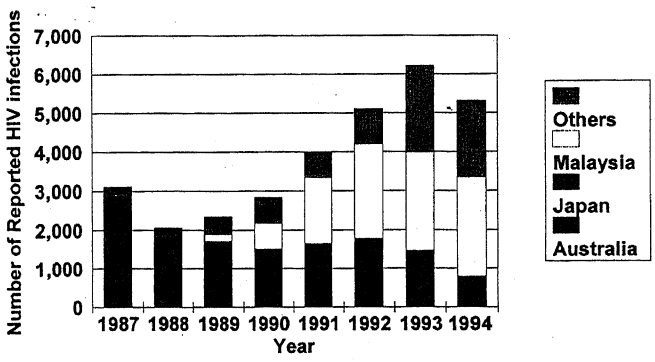


Figure 2. Reported incidence of HIV, Western Pacific Region 1987–1994.

Table 1. Cumulative HIV diagnoses and AIDS cases by country, sex and rates per 100,000 population, Western Pacific Region

Country/area	AIDS Males	AIDS Females	AIDS Unkn.	AIDS Total	AIDS rate[1]	AIDS deaths	HIV[2] Males	HIV[2] Females	HIV[2] Unkn.	HIV[2] Total	HIV rate[1]
American Samoa	0	0	0	0	0.00	0	0	0	0	0	0
Australia	4894	180	0	5074	28.03	3517	15346	840	2089	18275	100.97
Brunei Darussalam	4	0	0	4	1.37	2	175	7	0	182	62.22
Cambodia	0	0	1	1	0.01	1	362	126	145	633	6.62
China	41	2	0	43	0.00	24	1426	124	0	1550	0.12
Cook Islands	0	0	0	0	0.00	0	0	0	0	0	0.00
Fiji	4	3	0	7	0.90	7	15	6	0	21	2.71
French Polynesia	25	5	12	42	19.39	34	100	37	1	138	63.73
Guam	23	1	0	24	16.93	16	55	9	0	64	45.15
Hong Kong	108	7	0	115	1.98	70	442	38	10	490	8.43
Japan	769	41	0	810	0.65	483	2717	764	0	3481	2.77
Kiribati	0	0	0	0	0.00	0	0	0	2	2	2.55
Lao P D R	11	1	2	14	0.30	10	31	38	11	80	1.71
Macao	7	1	0	8	1.92	7	23	64	1	88	21.08
Malaysia	101	14	0	115	0.59	79	9653	380	15	10048	51.57
Mariana Islands, N	0	0	6	6	10.44	5	9	0	1	10	17.39
Marshall Islands	4	2	0	6	11.35	2	5	3	0	8	15.14
Micronesia, Fed Sts	2	0	0	2	1.75	1	2	0	0	2	1.75
Nauru	0	0	0	0	0.00	0	0	0	0	0	0.00
New Caledonia	33	4	0	37	19.92	20	90	21	0	111	59.76
New Zealand	432	19	0	451	12.73	359	867	69	17	953	26.91
Niue	0	0	0	0	0.00	0	0	0	0	0	0.00
Palau	1	0	0	1	5.83	1	1	0	0	1	5.83
Papua New Guinea	45	42	0	87	2.13	42	124	111	1	236	5.78
Philippines	104	62	0	166	0.25	95	290	260	7	557	0.85
Republic of Korea	20	5	0	25	0.06	22	343	43	0	386	0.87
Samoa	1	0	0	1	0.62	1	1	0	0	1	0.62
Singapore	83	6	0	89	3.04	61	230	22	0	252	8.59
Solomon Islands	0	0	0	0	0.00	0	0	0	0	0	0.00
Tokelau	0	0	0	0	0.00	0	0	0	0	0	0.00
Tonga	5	0	0	5	5.07	5	5	1	0	6	6.09
Tuvalu	0	0	0	0	0.00	0	0	0	0	0	0.00
Vanuatu	0	0	0	0	0.00	0	0	0	0	0	0.00
Viet Nam	197	18	0	215	0.30	64	1557	222	30	1809	2.51
Wallis and Futuna	1	0	0	1	7.14	0	2	0	0	2	14.29
Totals	6915	413	21	7349	0.45	4928	33871	3185	2330	39386	2.42

Data from reports received at WPRO as of 1 February 1995.
(1) Rate per 100,000 population based on WHO/WPR/HIN/Rev/June 1994.
(2) Totals include AIDS cases.

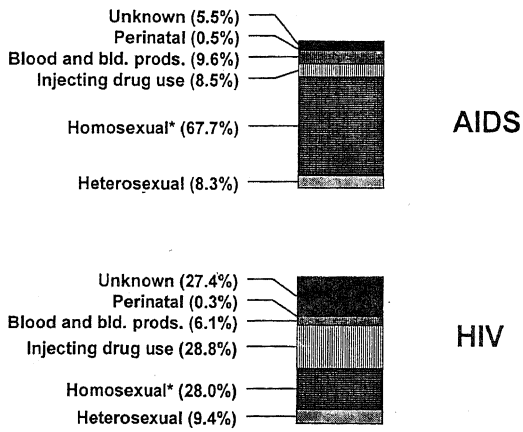


Figure 3. Transmission categories, AIDS and HIV, Western Pacific Region, cumulative to 1 February 1995. Note: HIV cases of unknown origin excluded. *Includes bisexual transmission.

are detected through screening procedures such as one would have prior to donating blood. Blood donors, however, represent only a small fraction of the population[7]. One can assume that there are others who are infected but who do not belong to the blood donor population.

To plan, develop and implement programmes, knowledge of the actual magnitude of HIV infection is necessary. Some countries have made estimates. These are often based on the number of reported cases, observed HIV prevalence rates, and the assumed size of the populations at risk. Australia, for example, estimates an HIV prevalence of 20,000–25,000, which is quite close to the reported figure. The Philippines, however, estimate the number of HIV-infected people at 35,000–50,000 compared to a reported 557 HIV infections.

Most countries in the Western Pacific Region have few reported cases, which makes reliable national and regional estimates difficult. Keeping these provisos in mind, WHO estimates that between 50,000 and 100,000 people are already infected in the region.

HIV and AIDS transmission in countries of the Western Pacific Region

The HIV epidemic in the Western Pacific Region has been characterized by wide variations between countries in the predominant modes of transmission, reflecting their great social, cultural and economic diversity and the measures taken to control the spread of the virus. Figure 3 shows the distribution of transmission categories for AIDS cases and HIV infections. It is important to note that AIDS figures represent the status of HIV approximately 10 years before, so that differences in rates in the modes of transmission indicate shifts in the way HIV is spread from one period to another. Close to 68% of AIDS cases in this region have been acquired through homosexual or bisexual contact. Injecting drug use and heterosexual transmission were responsible for 8.5% and 8.3% of AIDS cases, respectively. Note the decreased proportion of HIV infection among homosexuals and bisexuals and the increased proportion among heterosexuals and injecting drug users. This indicates the greater recent importance of the latter two risk categories in the transmission of HIV.

Figure 4 shows that homosexual transmission remains an important mode of transmission in many countries, including Australia, French Polynesia, Hong Kong, and New Zealand. In Australia, however, the number of homosexual men diagnosed with HIV is gradually declining (Figure 5). Another important feature of the Australian epidemic has been the low proportion of HIV infections among injecting drug users, which may explain why Australia has not seen an epidemic among heterosexuals.

Although the HIV epidemic in Japan was initially confined to people who had received blood products or blood transfusions prior to the screening of blood for HIV, during the last few years heterosexual transmission has been increasing (Figure 6).

Heterosexual transmission of HIV is predominant in many Western Pacific countries (Figure 7). In the Philippines, more than 50% of the reported infections have been acquired by this mode of transmission. In the Republic of Korea, eight out of ten AIDS cases and almost 70% of HIV diagnoses have resulted from heterosexual contact. Cambodia, Lao People's Democratic Republic and Papua New Guinea show heterosexual transmission of HIV in almost all cases where exposure was known.

In a few countries the role of injecting drug use in the transmission of HIV is becoming increasingly prominent. In Malaysia, 79% of the reported HIV infections

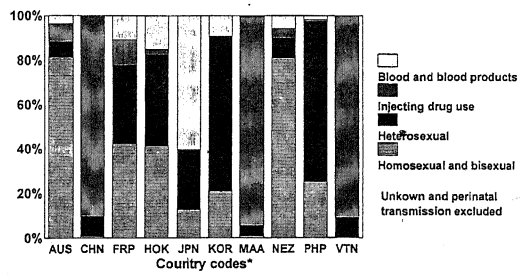


Figure 4. Transmission categories, HIV cases, selected countries, cumulative to 1 February 1995.

*AUS – Australia; CHN – China; FRP – French Polynesia; HOK – Hong Kong; JPN – Japan; KOR – Republic of Korea; MAA – Malaysia; NEZ – New Zealand; PHL – Philippines; VTN Viet Nam.

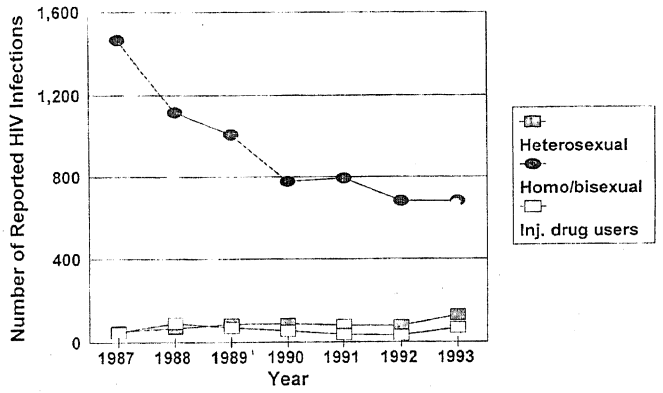


Figure 5. Trends in HIV infections, Australia, 1987–93.

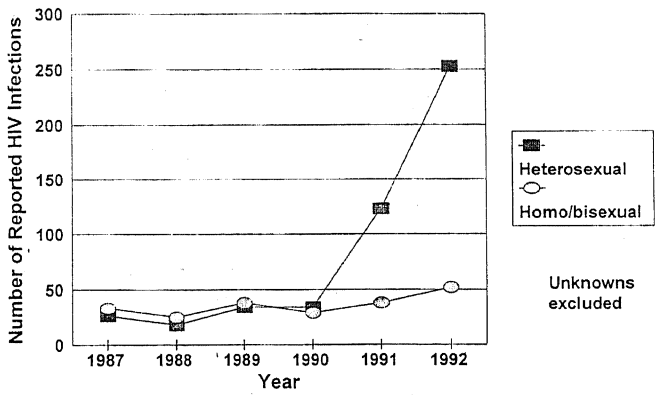


Figure 6. Trends in HIV infection, Japan, 1987–92.

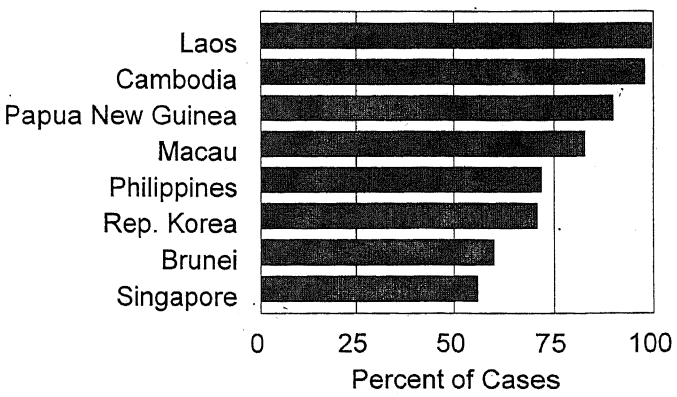


Figure 7. Heterosexual transmission, percentage when risks are known.

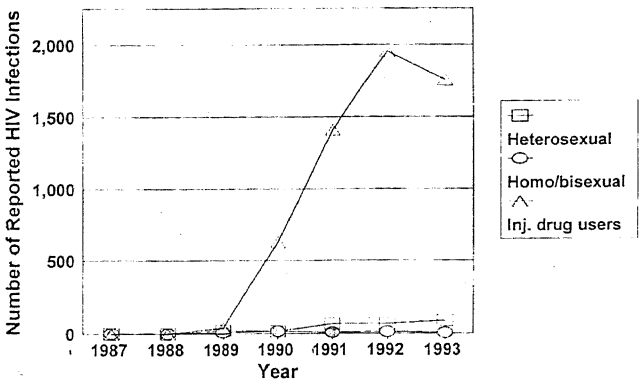


Figure 8. Trends in HIV infections Malaysia, 1987–93.

were attributed to injecting drug use. This number has increased dramatically since 1989 (Figure 8). More recently, Viet Nam has reported increasing number of HIV infections, over 80% of which is attributed to injecting drug use. China has also reported a significant number of HIV-infected injecting drug users, who account for 69% of all HIV transmission.

HIV has recently spread in Cambodia, reporting significant numbers of HIV infections for the first time in 1992 and 1993. Although only 1 AIDS case has yet been reported in Cambodia, more than 600 people have been diagnosed with HIV infection.

Risk factors for HIV transmission in the Western Pacific Region

Monitoring HIV infection in high-risk groups such as prostitutes, people with sexually transmitted diseases, injecting drug users, and homosexual or bisexual men is an essential part of HIV surveillance programmes. However, the results have to be carefully interpreted as survey methods differ. Unlinked anonymous testing is the method recommended for seroprevalence surveys and sentinel surveillance. In practice, voluntary confidential testing, compulsory testing of blood donors, and in some cases mandatory testing, are also used. As the test methods are not always specified and selection bias cannot be excluded, the results reported should be taken only as indicative of the situation in the areas surveyed. For example, sentinel surveillance of HIV infection in sexually transmitted disease clinics in Australia[8–10] which involve relatively stable population sources, performed at regular intervals and using unlinked anonymous testing, will yield far more reliable data than the irregular surveillance of STD patients in Cambodia.

Prostitution

HIV has recently spread rapidly among prostitutes in countries such as Thailand and India, which are close to the Western Pacific Region. With the exception of Cambodia, available data indicate that the spread of HIV among prostitutes has not yet occurred to the same extent in the region, even though sex industries catering to both local and foreign clients are well established in many countries.

The highest prevalence of HIV among prostitutes to date has been found in Cambodia. In one area, Toul Kork, 18 of the 127 (13.2%) women tested were HIV-infected. In a coastal area in the South, an unrepresentative sample of over 200 subjects showed[11] a prevalence of 38%. In Malaysia HIV prevalence among prostitutes has increased gradually, from zero in 1988 to 1.3% in 1991. In the mid-1980s, the prevalence of HIV among prostitutes tested in the Philippines was between 1.0 and 2.5%. Recent surveys, however, in Baguio and Cebu in the Philippines found no infection among more than 800 prostitutes tested, although testing bias could not be excluded. More recently, a survey in Metro Manila suggests a prevalence of 0.1–1.0% in this population group[12].

In the Republic of Korea, 411,125 tests performed between 1985 and 1991 revealed only 11 positive cases. No infection has been documented recently among prostitutes in Australia, Japan or Singapore.

Injecting drug use

The problem of HIV transmission through injecting drug use is escalating in Malaysia, where prevalence among injecting drug users rose from 0.1% in 1988 to 6.9% in 1992. In Viet Nam in early 1993, 92 out of 1061 (8.7%) of injecting drug users were found to be infected, most of whom came from Ho Chi Minh City. This has since increased to a cumulative percentage of 81% of 1809 reported HIV infections as of 1 February 1995. In China, injecting drug users infected with HIV were concentrated in Yunnan Province. Recent data, however, have indicated infection in other provinces as well. In Australia, the prevalence of HIV among injecting drug users remains at less than 5%. This may be attributable to the extensive efforts made to educate and protect injecting drug users in Australia, particularly with needle and syringe exchange programmes.

Homosexual and bisexual activity

Surveys conducted among men who have sex with other men show a range of prevalence from a localized high of 30.3% in Sydney, Australia, to zero in the Philippines. Japan and Singapore show a prevalence of 2.0% and 3.6%, respectively. Although the situation has stabilized in countries like Australia and New Zealand, data are still lacking on this transmission category in such countries as Cambodia, China, the Lao People's Democratic Republic, Malaysia, the Republic of Korea and Viet Nam.

Use of blood products

Blood transfusion and the use of blood products for haemophiliacs account for 6.1% of HIV transmission in the Western Pacific Region. The majority of these cases have been reported by Australia and Japan and represent transmission prior to the introduction of screening of the blood supply in 1985.

Perinatal transmission

Surveys performed to date among pregnant women have found virtually no HIV infection. This may be the reason for the less than 60 reported HIV infections and nineteen AIDS cases transmitted from mother to child within the region.

Sexually transmitted diseases

Systematic surveillance among those suffering from sexually transmitted diseases has been carried out in Fiji, Hong Kong, Papua New Guinea and Singapore. These and other studies show low levels of infection. In Cambodia, however, a recent study found 4.2% of 72 patients were infected. For the remaining countries, the lack of data is a result of insufficient surveillance in this category.

Future trends

Any forecasts of the future magnitude of HIV infection in the Western Pacific Region are bound to include wide margins. This is because there are many determinants of the infection and their significance for the region has not been thoroughly assessed. Present estimates suggest a two- to four-fold increase in the prevalence of HIV infection for the Western Pacific Region by the year 2000. This would mean that 100,000 to 400,000 people would be infected. However, many factors unique to the region may affect these figures; for example, the high level of literacy in most countries in the region, the large numbers of people with access to radio and television, the wide circulation of newspapers, magazines and books, the well-developed health infrastructure, especially in respect of maternal and child health and family planning, and the relatively low prevalence of ulcerative sexually transmitted diseases may lead to a lower prevalence of HIV. However, other factors, such as proximity to areas of high prevalence of HIV infection and injecting drug use, large-scale migration, and complacency may have the opposite effect. These factors will vary in importance in different countries. An essential part of HIV surveillance and prevention programmes is to single out the risk factors for spread and to implement locally appropriate strategies.

1. WHO/WPR/HIN/REV/June 1994.
2. Global Programme on AIDS, The Current Global Situation of the HIV/AIDS Pandemic, Geneva, 3 January 1995.
3. Sarda, R. *et al.* The Epidemiology of HIV/AIDS in the Western Pacific Region: An Opportunity for Prevention, Tenth International Conference on AIDS, Oral Presentation, Yokohama, 7–12 August 1995.
4. AIDS in The Western Pacific Region, WHO/WPRO Publication, July 1993, 2.2:3.
5. WHO/WPR/GPA Database.
6. AIDS Surveillance Report, Western Pacific Region, WHO/WPRO Publication, Manila, January 1995, no. 4.
7. Sarda, R., AIDS and HIV in the Western Pacific Region and Priorities for Prevention and Control, KIHASA, Seoul, February 1994, 1:2.
8. Australian HIV Surveillance Report, October 1993, vol. 10, no. 4.
9. Australian HIV Surveillance Report, January 1994, vol. 10, no. 1.
10. Australian HIV Surveillance Report, vol. 10, no. 2.
11. Official Communication, WHO Representatives Office, Phnom Penh, October 1994.
12. Official Communication, WHO Representatives Office, Manila, December 1994.

Heterosexual risk behaviour in Asia: The implications for HIV/AIDS

Tim Brown and Werasit Sittitrai*

Program on Population, East West Center, Honolulu, Hawaii
*Thai Red Cross Society Program on AIDS, 1871 Rama IV Road, Bangkok 10330, Thailand

IN the early 1980s the belief that Asia might avoid a major HIV/AIDS epidemic was widespread. As the epidemic spread in the Western world and devastated Sub-Saharan Africa, Asia appeared strangely immune. By the end of 1987, Asian countries had reported only 252 AIDS cases to the World Health Organization[1], and when a case did appear in Asia, it was most often associated with Asian men who had sex with men who had sexual contact with foreigners or had travelled abroad or with haemophiliac's use of foreign blood products. Indigenous transmission in observed AIDS cases was rare.

To explain this low rate some speculated that Asians might have a natural resistance to HIV infection. Others said that traditional Asian morality and the value placed on families were protecting the continent from the epidemic. The region seemed to have little cause for concern, and most countries did not undertake any large-scale efforts to prevent or slow down the spread of HIV.

In early 1988, this belief in Asia's immunity to the AIDS epidemic was shaken as HIV prevalences in injecting drug users in Thailand increased rapidly from zero to 30% over an 8-month period[2]. At the same time, studies of female commercial sex workers in Bombay and Madras revealed[3] HIV infection levels of 5%. Prior to this time, some HIV testing had been done among populations vulnerable to HIV in Asian countries, including Thailand and India, but had rarely found infection levels exceeding 1% in any groups other than haemophiliacs and men having sex with men. During the next few years, both India and Thailand saw rates of HIV infection climb first in injecting drug users and commercial sex workers, then in male patients with sexually transmitted diseases (STDs), and finally in pregnant women. For example, by mid-1993 nationwide median HIV prevalences in STD patients and antenatal clinics in Thailand had grown to 8.0% and 1.4%, respectively[4]. In Bombay by 1993 STD clinics showed rates as high as 26% and rates among antenatal women were approaching 1% (ref. 3). Studies of risk factors done in India and Thailand found that most of the infections had been contracted through heterosexual intercourse. Although the rapid spread of HIV started among injecting drug users and commercial sex workers, by the early 1990s the epidemic had clearly established itself in the general population.

Because most other Asian countries have not conducted as extensive HIV testing as India and Thailand, their situation is more difficult to assess. In most countries of the region, the number of reported AIDS cases is small, as seen in Table 1, which shows the cumulative number of AIDS cases reported through the end of 1990 and through mid-1994. In many of the countries the HIV/AIDS epidemic has remained largely invisible, although the increasing AIDS numbers show that an epidemic is underway in most of them.

Several factors have contributed to this invisibility. The first is the nature of HIV itself. In the past many Asian countries have relied largely on reported AIDS cases for data, with only limited HIV testing. Because illness takes many years to develop, and the introduction of HIV in much of Asia occurred in the mid-1980s, case

Table 1. Number of AIDS cases reported to the World Health Organization by Asian countries through December 1990 and June 1994. Source: Global Programme on AIDS 1994

Country	December 1990	June 1994
Bangladesh	1	1
Bhutan	0	0
Cambodia	0	0
China	5	36
Democratic People's Republic of Korea	0	0
Hong Kong	27	99
India	52	713
Indonesia	9	49
Japan	294	713
Laos	0	14
Malaysia	15	107
Mongolia	0	0
Myanmar	0	261
Nepal	4	24
Philippines	37	136
Republic of Korea	7	19
Singapore	21	75
Sri Lanka	6	37
Thailand	69	5654
Vietnam	0	107

numbers remain low. In addition, reported numbers are lower than the actual number of cases because HIV disease manifests itself through other diseases, such as tuberculosis and pneumonia. As a consequence, AIDS is often misdiagnosed or goes unrecognized, especially when physicians are not expecting it. This, along with the social stigma of AIDS in most Asian societies, has resulted in underreporting of AIDS cases.

The second factor is the lack of extensive HIV surveillance systems similar to the sentinel seroprevalence system in Thailand[2]. Most countries have only recently begun to establish systematic prevalence testing on even a limited scale. Thus, the quality and quantity of the available HIV data are only now beginning to improve.

The third factor is the public reluctance in Asian societies to discuss or admit to the behaviours that transmit HIV: injecting drug use; homosexual intercourse; and extramarital or premarital commercial and casual sexual intercourse. In most Asian countries, the limited studies of injecting drug use, sexual partnering patterns and sexual practices have focused primarily on attitudes toward drug use and extramarital and premarital sex, not on actual behaviour. While the surveys generally find that publicly expressed attitudes toward sexual activities outside marriage are negative, men are often allowed great latitude to engage in them without sanctions. These negative attitudes and the public reluctance to discuss the issues give the false impression that the risk behaviours needed to sustain an epidemic do not exist in many Asian societies.

However, despite these limitations sufficient HIV and AIDS data do exist to reveal certain similarities and differences between Asian countries. For example, the rates of HIV spread within different countries vary significantly. Although the first AIDS cases appeared in both Thailand and the Philippines in 1984, by 1993 infection levels among commercial sex workers were much lower in the Philippines, 0.6%, compared to approximately 9% among Thai nonbrothel-based commercial sex workers[4,5]. Current data indicate that the rate of spread is quite high in some countries, for example, Cambodia, India, Myanmar and Thailand, but is substantially slower in others, such as China, Hong Kong, Japan, the Philippines and the Republic of Korea.

The relative importance of different transmission modes also varies substantially. In Japan, haemophilia patients who received blood products imported mostly from the United States account for 64% of all AIDS cases[6]. In China and Malaysia most AIDS cases cite injecting drugs as the route of exposure[7,8], but in Japan fewer than 1% of nonhaemophiliac AIDS cases are attributed to injecting drugs[5]. In Hong Kong, India, Indonesia, and Thailand most infections in reported AIDS cases have been through sexual intercourse[3,4,9,10]. However, in Hong Kong and Japan infection through homosexual intercourse is more common, while heterosexual intercourse is the more frequent transmission route in the other countries.

The HIV/AIDS epidemic in Asian countries is dynamic in nature, and the comparative contribution of the various transmission routes is changing with time. One striking similarity among the countries of Asia is that most are either seeing or moving toward predominantly heterosexual transmission of HIV. For example, in Hong Kong, while the overall ratio of homosexually transmitted to heterosexually transmitted HIV infections is 1.17 to 1.00, if one only considers those infected in 1993, the ratio is 1 to 2 in favour of heterosexual infections[9]. Japan has almost twice as many detected heterosexual HIV infections as homosexual infections, and new heterosexual AIDS cases have outnumbered homosexual AIDS cases since 1992 (ref. 6). In Singapore, a second wave of heterosexual infections is now following the initial wave of homosexual infections[11]. There is little doubt that in the long term the HIV/AIDS epidemic in Asia will be principally a heterosexual epidemic. Heterosexual risk behaviour is thus playing a key role in the epidemic.

The roles of commercial, casual and marital sex in the heterosexual epidemic

If the course and severity of the epidemic in Asia are to be anticipated and an understanding reached of the variation among countries, exploring the forms of heterosexual transmission and their relative contributions to the epidemic is important. Heterosexual transmission can occur in three settings: during commercial sex, during noncommercial casual sex outside marital relationships, and during marital coitus. Comparing the relative rates of transmission in these settings and examining the prevalence of each in Asian societies may help to clarify the situation. Before looking more closely at these factors, a brief examination of the Thai situation can provide valuable insights into the factors influencing a heterosexual epidemic.

The Thai HIV/AIDS epidemic has been one of the most intensively studied epidemics in history. Although the first Thai AIDS cases occurred in 1984, the epidemic did not begin to spread rapidly until late 1987 among injecting drug users. In 1989, comparatively early in the epidemic, Thailand established national surveillance for HIV that has allowed tracking of the epidemic's progress[12,13]. Based on these data, collected by the Ministry of Public Health and the Royal Thai Army, and on extensive epidemiological and behavioural studies, the patterns of the Thai epidemic are well understood. Five waves have been observed: (1) injecting drug users; (2) female commercial sex workers; (3) clients

of commercial sex workers; (4) wives and girlfriends of clients and (5) children born to infected mothers. (For more detailed material on the Thai epidemic, see refs. 2, 4).

Unlike HIV/AIDS epidemics in the West, where the contribution of heterosexual transmission has been low, the Thai epidemic has seen a rapid spread among heterosexuals, with infections among female commercial sex workers spreading particularly fast. For example, limited testing of sex workers in Chiangmai in northern Thailand in 1987 found no HIV, but by 1989 levels had surpassed 40% in brothel based sex workers[2]. This explosive growth in infection among brothel-based female commercial sex workers resulted both from the large numbers of clients they serviced and from the sex workers' high levels of other STDs, which enhance HIV transmission. At the same time, infection levels among indirect sex workers, working out of bars, nightclubs and restaurants grew more slowly because both their client numbers and STD levels were lower. For example, in mid-1990, when HIV prevalence among brothel-based sex workers was 14% nationwide, it was less than 3% among indirect sex workers.

Despite the rapid growth in infection among sex workers, the bulk of the country's present HIV infections resulted from the third and fourth waves: clients of sex workers and their female sexual partners. In 1990, the Survey of Partner Relations found that 23.4% of Thai men between 15 and 49 years of age had engaged in commercial sex in the last year[14]. This large pool of susceptible males, in conjunction with high rates of female sex worker to client transmission[15], contributed to a rapid rise in new male infections between 1989 and 1991 (ref. 16). The rate of new male infections has now slowed substantially because of an aggressive national program that has increased condom use, brought about a substantial drop in STD levels, and reduced the number of men visiting sex workers, but the third wave has left behind close to half a million infected males.

The fourth wave, wives and girlfriends of clients of commercial sex workers, proceeds at a slower but relentless pace. Recently completed studies have found that transmission from infected husband to wife occurs at a much lower rate than from sex workers to clients[17]. This probably results because rates of STD co-infection are lower in both marital and noncommercial casual sexual contacts than in commercial contacts. However, because married couples and noncommercial sexual partners rarely use condoms in Thailand, in the absence of intervention most infected men will eventually infect their wives or girlfriends, although it may take several years. Thus, female infections, which currently number about one quarter of a million, lag substantially behind male infections. This is in contrast to the situation in Africa, where male and female infection levels are roughly equal, and reflects the important role played by commercial sex in the early dynamics of the Thai epidemic. The extent to which unmarried Thai women are at risk depends largely on how males' sexual networks link casual and commercial sex. Because women married to infected men outnumber sex workers, most infected women in Thailand contracted their infections from their husbands, despite the higher HIV transmission rates in commercial sex.

Other countries can learn several important lessons from the Thai experience. If a country has a sex worker population with high client numbers and STD levels, then the growth of HIV can be explosive. Given that HIV had been introduced to Thailand as early as 1984, but that infections among commercial sex workers were rarely found until 1988, this indicates that a substantial delay can occur between HIV introduction and the beginning of this explosive growth. The absence of detected HIV infections in commercial sex worker populations today should not be misconstrued as the absence of potential for a rapid spread of HIV. The much lower seroprevalence among indirect workers compared to brothel-based workers shows that the frequency with which commercial sex workers have clients has a significant effect on the rate of infection on a national scale. Thus, the relative mix of brothel-based commercial sex with other forms of commercial sex may provide some indication of the overall vulnerability to the rapid spread of HIV. Finally, the rate of HIV transmission in commercial contacts and in marital or noncommercial casual contacts can vary greatly, perhaps by an order of magnitude or more. Thus, the heterosexual epidemic may proceed much more slowly in countries where noncommercial casual sex is more common than commercial sex.

As noted earlier, the predominantly heterosexual nature of the AIDS epidemic is becoming apparent in most Asian countries. Based on the analysis of the Thai epidemic presented here, the rate of growth and ultimate magnitude of the HIV/AIDS epidemic in each Asian country will be determined largely by: the prevalence of heterosexual commercial sex; the overlap between commercial, casual and marital sex; and the levels of other sexually transmitted diseases. To assess whether a particular country or population is likely to face major HIV problems, these are the factors that must be examined. The remainder of this paper will, therefore, focus on the levels of heterosexual risk in the countries of Asia to assess the potential for the spread of HIV in the region as a whole.

Heterosexual risk behaviour in Asia – commercial sex

As India's and Thailand's experiences have shown, commercial sex can play a critical role in HIV transmission in Asia. While commercial sex is certainly not unique to Asia, it may have more of a role in epidemics in the region than in other parts of the world.

Commercial sex, despite frequent public condemnation, has a long history of tacit acceptance in many Asian societies. As a consequence, it is widespread, and sex workers often have large numbers of clients. For the human immunodeficiency virus, these high client numbers offer two obvious advantages: (1) commercial sex workers have higher than normal rates of infection and reinfection with other sexually transmitted diseases, which enhance HIV transmission; and (2) the larger numbers of sexual partners offer increased opportunities for the virus to spread from client to sex worker and from sex worker to client. These factors have driven the rapid growth of HIV seen in Cambodia, India, Myanmar and Thailand.

Efforts at HIV prevention during commercial sex are often hampered by the comparative lack of control sex workers have over sexual transactions in Asian societies, where control usually rests with the client. Thus, even if a sex worker wants to take precautions against disease transmission, for example, by using condoms, this is not her choice to make. This lack of control greatly increases the workers' vulnerability to HIV and other STDs.

The roots of commercial sex in Asia are both cultural and economic. Culturally, widespread gender inequality in sexual matters is the most important factor contributing to the demand for commercial sexual services. Many Asian societies have very different expectations from men and women in terms of sexual behaviour. Women are almost universally expected to be virgins at marriage and not to engage in extramarital sexual activities. For men, by contrast, despite publicly expressed attitudes opposing premarital and extramarital sex, strong societal sanctions against such practices are lacking and in some places men are expected to be sexually experienced at the time of marriage or people believe that men need sexual variety. This creates an imbalance between male sexual demand and female sexual supply. While many men seek sexual contact outside of marriage, either premaritally or extramaritally, few women are available. This creates the high demand for female commercial sexual services.

Growing economic prosperity and economic disparity between urban and rural areas are also fuelling the growth of commercial sex in the region. While the countries of Asia as a whole are experiencing phenomenal economic growth, the resulting benefits are not equally distributed between urban and rural areas. While improving economic conditions drive an increasing consumerism, permeating the rural villages, where most Asians live, the resulting employment opportunities are generally concentrated in urban areas. This increases the urban/rural economic disparity, leaving villagers without the means to pay for either the basic necessities of life or for the consumer items and luxuries that are becoming national norms. Lacking the education or training necessary for many of the newly created jobs, many young female villagers turn to commercial sex as an alternative for supporting their families. The steady flow of men and women into urban areas for work also removes them from their traditional social environments and disrupts social norms, thereby further increasing the demand for and acceptability of commercial sex.

Economic disparity is growing not only between urban and rural areas, but also between countries in the region. This has fuelled an active international sex trade encompassing both sexual tourism and commerce in female sex workers. Men visiting commercial sex workers while travelling is commonly reported throughout Asia (see, for example, ref. 18 on Japanese men in Thailand and ref. 19 on Korean men). As education, improving economic conditions, and the fear of HIV reduce the supply of sex workers in Thailand, the deficit has been made up by sex workers with lower education levels and less knowledge of the AIDS epidemic imported from Myanmar and China. This is a regionwide phenomenon: Nepalese sex workers often go to India for work, Vietnamese sex workers are frequently found in Cambodia, and Thai sex workers are regularly imported into Japan. Should any of these women test positive for HIV during their foreign employment, they are inevitably returned to their country of origin. As a consequence, many of the earliest female AIDS cases in Asia are women who formerly worked for commercial sex in other countries and may have contracted HIV there.

The tacit acceptance of commercial sex and the high male sexual demand, combined with poverty and increasing consumerism, make commercial sex work an economically attractive alternative for women, especially those with little education or from poor families, and the supply/demand imbalance guarantees that female commercial sex workers will have a steady flow of clients, making it a viable profession.

The quiet acceptance of commercial sex is fostered by strong historical traditions in many Asian cultures. In a number of countries, laws against commercial sex are historically recent. Given the time often needed to effect changes in cultural beliefs and attitudes, in many countries the legal sanctions against commercial sex may be in direct conflict with an underlying cultural acceptance of the practice, resulting in lax enforcement of these laws. For example, China has a long history of both government and privately run commercial sex, perhaps starting as early as the seventh century BC[20]. It was only when the communists came to power in 1949 that commercial sex was made illegal. Despite legal restrictions, the recent rapid economic expansion, especially in southern China, has fuelled the growth of commercial sex and arrests of sex workers and clients have been increasing. In the Ayuddhya period in Thailand (1350–1767), commercial sex was both legal

and subject to government tax, becoming illegal only in the 1950s (ref. 21). In India, the Hindu tradition of dedicating girls (*Devadasi*) to gods and goddesses lowered the stigma against these women engaging in commercial sex work and provided a cultural basis for accepting commercial sex[22–24].

The result of these cultural, traditional and economic factors is that every country in Asia has a commercial sex industry, although in most places it is technically illegal. Because the number of clients is high and drawn from all strata of society, commercial sex occurs in a number of settings and forms to cater to the diversity and economic resources of the clientele: brothels, tea houses, coffee shops, street walkers, barber shops, massage parlours, and so on. The relative mix of these settings varies substantially from country to country. In most countries, the extent of the sexual service industry has been poorly determined, largely because of its illegality and the resulting invisibility of the sites, public reluctance to discuss the issue openly, and the protection afforded from official scrutiny by its economic profitability. While investigators have conducted numerous studies of commercial sex workers in Asia, these have focused primarily on the women's health and working conditions and very few have examined the total number of sex workers or establishments. Similarly, studies exploring the extent and practices of men who visit commercial sex workers are a recent development, prompted largely by the HIV/AIDS epidemic. Many of these studies, for example, in Malaysia, the Philippines, Singapore and Thailand were part of the effort by the World Health Organization's Global Programme on AIDS to encourage and support studies of sexual risk behaviour worldwide[25]. In several cases, these studies have proven extremely valuable in making policy-makers aware of the potential problems their countries face.

While current data on clients are limited and not always strictly comparable because of variations in sampling, population subgroup, underreporting differences, and information gathered, the highest rates of commercial sex have been reported in southeast Asian countries and the lowest in east Asian countries. Nevertheless in each country that has examined the situation, the fraction of the total male population visiting sex workers is not insubstantial, thereby creating the potential for rapid spread of HIV. For example, in Thailand in 1990 the Survey of Partner Relations found that 23.4% of males had visited a sex worker within the last 12 months[14]. This number had dropped to 10.1% by 1993. In Phnom Penh, Cambodia, a survey in 1993 found that 41% of sexually experienced males had visited a sex worker within the last year[26]. A survey in Manipur state, India, found that 19.1% of married men had visited a commercial sex worker in their lifetimes[27]. East Asian numbers appear somewhat lower, but researchers suspect underreporting. For example, in a survey conducted by the Hong Kong AIDS Foundation in 1992, 5.5% of male respondents had visited sex workers in the last 12 months[28]. An earlier survey in Hong Kong had found that 22% of males had visited a sex worker by age 26 (ref. 29). Similar numbers have been noted in Singapore. In a sample of international travellers in Korea, 24% reported having paid for sex in the last 5 years[19].

Heterosexual risk behaviour in Asia – non-commercial casual sex – premarital and extramarital sex

Needle sharing and commercial sex have often provided the initial high-grade fuel for Asian HIV/AIDS epidemics, resulting in infection of a large fraction of injecting drug users, female commercial sex workers, and clients of sex workers. From the clients HIV spreads to their girlfriends and wives. Unless condom use among married couples is common as a contraceptive, husbands who are infected through commercial sex will almost invariably transmit the infection to their wives. These women are at high risk even though they are engaged in 'low-risk' behaviour, thereby illustrating the danger of not considering the risk of all potential avenues of heterosexual HIV transmission. Thus, assuming that all infected males eventually marry, the size of the female population at high risk of HIV infection is the same as the population of male clients. If one includes extramarital and premarital sexual contacts, and the possibility that males may have multiple female partners over a period of time, the female population at risk becomes larger than the male population at risk.

While comparatively few studies have looked at sexual networks in the region, many have inquired about levels of extramarital and premarital sex. The findings raise serious concerns throughout Asia. In general, the studies confirm the difference in levels of male and female extramarital and premarital sex. In Japan, for example, a 1990 survey found that 19% of urban married males reported extramarital intercourse in the last 12 months compared to only 8% of women[30]. A study in Eastern Shan State in Myanmar found that 40.3% of married men, but only 5.5% of married women, reported extramarital intercourse during their lifetime[31]. In Manipur state, India, 27.8% of unmarried men and 11.3% of unmarried women reported having had intercourse[27]. In Nepal 24.1% of males and 14.9% of females reported premarital sex. The investigators observed a strong association of seasonal migration with premarital and extramarital sex in this case[32,33]. Similar high levels of extramarital and premarital sex have been found in many other countries, even when gender differences are not reported. For example, in Malaysia 33% of sexually experienced men and 11% of married men reported casual sexual contacts in the last year[8]. In

Korea, 46% of males who travelled internationally reported extramarital sex in the last 5 years[19].

There are some indications that male and female behaviour may be changing among adolescents in some parts of Asia, with closer equality of male and female premarital sex. For example, in a survey of final-year Japanese high school students, 20.5% of males and 22.9% of females reported sexual experience[34]. A similar study of students in Hong Kong found that 21% of males and 19% of females reported sexual experience[35]. How this translates into changes in partnering patterns and the relative mix of commercial and casual sex is not known.

The final factor that will influence the growth of the heterosexual epidemic is the linkage between commercial and casual sex. Studies of sexual networks in Thailand have found that most of the men visiting commercial sex workers often have one or more girlfriends prior to marriage, which may place their casual partners at risk[36,37]. Studies in Hong Kong and Singapore, in contrast, have shown less overlap between commercial and casual sexual relations, with most males having primarily commercial or casual contacts, but usually not both, prior to marriage. The total risk to women having casual sex with clients of sex workers may thus vary greatly from country to country. However, it should be remembered that HIV risk in a sexual network is not confined to simultaneous overlap of casual and commercial partners but is cumulative. If a single man engages in commercial sex while young and becomes infected, then all subsequent partners, including his future wife, are at risk.

Implications for the HIV/AIDS epidemic

Although most Asian countries report low numbers of AIDS cases and detected HIV infections, the trend is clear: heterosexual sex is becoming the dominant contributor to the growth of the HIV/AIDS epidemic in Asia. What does this mean for the future progression of the Asian epidemic, for intervention, and for future research needs?

Despite the limited availability of behavioural and epidemiological data, an explanation of the variation between countries is possible. Those countries seeing a rapid spread of HIV, for example, Cambodia, India and Thailand, are countries where commercial sex is common and a large percentage of the male population are clients of commercial sex workers. The resulting large number of nightly clients for commercial sex workers combines with elevated levels of other STDs to produce extremely high HIV transmission rates. The result is the frighteningly rapid growth of the epidemic. Those countries with lower rates of growth, for example, Hong Kong, Japan and the Philippines, are those where casual sex is as common or more common than commercial sex (but where commercial sex services still exist, although to a lesser extent) with lower overall levels of heterosexual risk. The epidemic has spread more gradually in these countries, but even within these countries, commercial sex presents greater risk for contracting HIV, and most early male heterosexual AIDS cases are associated with it. Thus, heterosexual risk in commercial settings is becoming the dominant factor in reported AIDS cases.

Casual sex has played less of a role in Asian AIDS epidemics, both because it occurs less often than in the West or in Africa, and because the transmission rates associated with it are lower. The female/male imbalance in sexual activity in Asia has slowed the spread into the unmarried female population. This is not to say that interventions can ignore casual sex, especially in the light of recent trends toward increasing casual sexual activity among females of the younger generation at precisely those ages where most sexually transmitted diseases, including HIV, are concentrated. Casual sex also deserves serious attention because the low perception of associated HIV risk leads to low levels of condom use. Even in countries such as Thailand, where condom use in commercial sex has risen dramatically, it remains quite low in casual sex, despite a large reservoir of currently infected males.

However, while most male infections in Asia over the next several years will occur in commercial settings, most female infections will happen in marital settings. Because comparatively few Asian females engage in casual sexual contacts, most of their HIV risk accrues from the sexual practices, current or past, of their mates. These cases will grow more slowly because of low HIV transmission rates in the absence of STDs, but they will grow inexorably unless interventions can be developed to address the risk posed to Asian women by their spouses' premarital and extramarital sexual activities.

Although existing behavioural data are limited and much research remains to be done, some subregional patterns are becoming apparent. Most countries of southeast Asia share a common pattern of tolerance toward commercial sex and frequent use of commercial sex services by males. This leaves these areas vulnerable to the same explosive growth seen in Thailand, and indeed, this is now being seen in Cambodia, Myanmar, and perhaps Vietnam. In east Asia, data are more limited, but studies indicate that commercial and casual sex occur at lower levels than in southeast Asia, with less overlap between casual and commercial sex. Nonetheless, a substantial portion of the male population has made use of commercial sex services in their lifetimes. These countries still have large enough populations of clients for the heterosexual epidemic to grow quite large. It will just take more time to get there. South Asia is somewhat of an enigma with fewer behavioural studies done, as well as substantial

variations in behaviour and practices even within countries, for example, in India. This has resulted in geographically uneven growth of the epidemic in India and makes assessing the country as a whole more difficult. No behavioural risk or epidemiological data are available for countries such as Bangladesh and Pakistan. However, the absence of data does not imply an absence of risk, and the case of India shows the potential for HIV spread in the south Asia subregion.

The heterosexual nature of the AIDS epidemic in Asia has consequences for the ultimate size of Asian epidemics. In one sense, epidemics in men having sex with men and in injecting drug users are self-limiting, because they constitute a comparatively small fraction of the total population. For example, in most places injecting drug users make up less than 1% of the total population and men having sex with men may represent 2–3%. Heterosexuals, however, make up the bulk of any population; thus, the group of susceptible people is much larger. The risk for individuals is also raised depending upon their age. For example, while in Hong Kong and the Philippines only a few per cent of males visit commercial sex workers, these men are concentrated in the 15–30 age range, meaning that the risk for individuals in this age range is substantially elevated.

This analysis has several implications for intervention programmes. Urgent intervention to reduce HIV transmission in commercial sex is essential. The national programme in Thailand has demonstrated that condom promotion, aggressive STD control, and mass media campaigns can radically reduce the rates of disease transmission in commercial sex[38]. These efforts have succeeded largely because authorities worked with brothel owners and sex workers to increase condom use and improve STD treatment, while simultaneously using mass media and workplace interventions to reach clients, rather than attempting to eliminate commercial sex. Given the cultural and economic forces sustaining the commercial sex industry, eradication in the short term is an idealistic, not a realistic, option. If one merely removes the supply of sexual services without altering the demand, the demand will express itself elsewhere. Commercial sex will be driven underground and intervention will become impossible. Such efforts will have the net result of promoting the spread of HIV. For now, it is better to focus on decreasing the HIV risk associated with commercial sexual contact through condom promotion and STD control.

From both the public health and human rights viewpoints there are clear reasons to discourage commercial sex, but this is a long-term, not a short-term strategy. Efforts should begin now to change the norms and practices of young people regarding commercial sex before their sexual habits have formed. Mass media approaches can be used to reach the public and show the negative aspects and risks associated with commercial sex. However, such efforts must be handled with care so as not to stigmatize further sex workers. Other long-term approaches include working on the roots of the commercial sex industry: poverty and social displacement. Development policies can increase employment and educational opportunities for young women, especially in rural areas, so they have fewer incentives to engage in commercial sex work. This might be done by providing scholarships for young, rural women to allow them to continue their education beyond primary school. In developing new industrial zones, consideration should be given to moving entire families rather than asking husbands to leave their wives and children behind when they leave to find work. This will help by providing a more normal social environment, thereby reducing the demand for sexual services. Finally, efforts to improve women's status and their control in sexual situations should continue, because the roots of Asia's HIV problem lie in the inequality between the genders.

Intervention in the casual and marital sex arenas are also needed. At present condom use in casual sexual contacts is almost nonexistent in most of Asia. As the heterosexual spread of HIV infection continues, this places large numbers of people at risk for contracting HIV. Focusing on this is particularly important in countries with strong sexual network linkages between casual and commercial sex, for example, in southeast Asia. General efforts to educate the public should stress that all vaginal sex carries a risk of HIV transmission, focusing on the act rather than on the partner. Too often in the past the emphasis has been on commercial sex in education messages, not on vaginal transmission, leaving many with the impression that it was who you had as a partner, not what you did, that placed you at risk. Education efforts should teach proper use of condoms and negotiating skills for sexual situations. Another essential component of programmes to address casual sex should be the reduction of barriers to access to condoms and STD treatment for single men and women. In many places, for example, Indonesia[39], unmarried people face significant social and structural barriers to condom access or STD treatment.

For women in Asia, the greatest risk of HIV comes from regular sexual partners, primarily husbands. In marital settings, approaches to increase sexual communication between husbands and wives should be explored. Married Asian women need to understand the risks they face. This will only happen if husband and wife discuss sexual matters and come to a mutual understanding. Part of these efforts will involve education to make women aware of the levels of premarital and extramarital activity in their countries, and part will be training in skills and building mechanisms of peer support so that women are in a

position to influence their husbands' behaviour. In addition, governments could consider promoting voluntary premarital or pre-pregnancy testing and counselling. If properly executed, such efforts might encourage men to reduce their premarital and extramarital risks, while providing women with the knowledge they need to protect themselves.

Finally, an examination of the existing situation emphasizes the need for expanding knowledge about sexual risk in Asian countries. While studies have been done in many countries, they are of varying quality and generalizability. Only a handful are scientifically executed nationwide samples that can be said to represent adequately the behaviours in the country. In many cases, studies have been weakened by high levels of nonresponse. In most Asian countries, little or no information is available on sexual behaviour in rural areas, where most people live. In other countries no representative sexual behaviour studies have been done at all. If national policies are to be based on a realistic appraisal of the level of HIV risk in a country, and if the effectiveness of intervention programmes is to be evaluated, regular, consistent and methodologically sound surveys need to be conducted at regular intervals. Given the importance of commercial sex in the Asian HIV/AIDS epidemics, such studies should explicitly gather information on both casual and commercial sex. Behavioural surveillance should be an essential component of every national AIDS programme, with national surveys executed every 2 or 3 years to track trends and changes. Two such studies done 3 years apart in Thailand found substantial increases in condom use along with a significant decline in visits to commercial sex workers, which indicates that the national programme is achieving its goals. Such regular behavioural surveillance should be coupled with intensive qualitative and quantitative studies of the forms and prevalence of commercial sex and the levels of HIV and STD among commercial sex workers.

A common misconception has been that because most countries in Asia have seen fewer AIDS cases than Thailand, they do not have sufficient levels of risk behaviour to support a widespread, heterosexual HIV epidemic. Recent shifts in epidemiological patterns and behavioural studies done in the region have shown this to be untrue. While the epidemic will vary in terms of its growth rate and ultimate size in different countries of the region, each of them has sufficient heterosexual risk behaviour to support a large-scale HIV/AIDS epidemic. Just because they have not experienced a large epidemic by 1994 does not mean that they are immune to one in the future, but it does present them with a unique opportunity to intervene early and prevent much death and suffering. An immediate investment in national programmes to reduce risk in commercial settings coupled with a strong education programme aimed at the general population will reap immense benefits in averted costs and social destruction. Immediate action is essential. Every month of delay only adds to the epidemic's eventual social and economic cost.

1. Global Programme on AIDS, AIDS Cases Reported to WHO by Continent/Year, Geneva, Switzerland, 1994.
2. Weniger, B., Limpakarnjanarat, K., Ungchusak, K., Thanprasertsuk, S., Choopanya, K., Vanichseni, S., Uneklabh, T., Thongcharoen, P. and Wasi, C., The epidemiology of HIV infection and AIDS in Thailand, *AIDS*, **5** (suppl. 2), S71–S85 (errata corrected in 1993, **7**(1): following 147).
3. Jain, M. K., John, T. J. and Keusch, G. T., Epidemiology of HIV and AIDS in India, *AIDS*, 1994, **8** (suppl. 2), S61–S75.
4. Brown, T., Sittitrai, W., Vanichseni, S. and Thisyakorn, U., The recent epidemiology of HIV and AIDS in Thailand, *AIDS*, 1994, **8** (suppl 2), S131–S141.
5. Tan, M. L. and Dayrit, M. M., HIV/AIDS in the Philippines, *AIDS*, 1994, **8** (suppl. 2), S125–S130.
6. Kitamura, T., Summary of the epidemiology of HIV/AIDS in Japan, *AIDS*, 1994, **8** (suppl 2), S95–S97.
7. Xinhua, S., Junhua, N. and Qili, G., AIDS and HIV infection in China, *AIDS*, 1994, **8** (suppl 2), S55–S59.
8. Singh, J., Che'Rus, S., Chong, S., Chong, Y. K., and Crofts, N., AIDS in Malaysia, *AIDS*, 1994, **8** (suppl. 2), S99–S103.
9. Lee, S. S., Lim, W. L., Lee, S. H. and Chan-Fung, M. F. C., Epidemiology of HIV Infection in Hong Kong Paper presented at the Tenth International Conference on AIDS and STDs, Yokohama, Japan, August 1994 (abstract PC0070), 1994.
10. Jalal, F., Abednego, H. M., Sadjimin, T., and Linnan, M. J., HIV and AIDS in Indonesia, *AIDS*, 1994, **8** (suppl. 2), S91–S94.
11. Oh, H. M. L. and Chew, S. K., HIV infection in Singapore, 1985-1993, Paper presented at the Tenth International Conference on AIDS and STDs, Yokohama, Japan, August 1994 (abstract PC0071).
12. Division of Epidemiology., First sentinel surveillance June 1989. *Weekly Epidemiol. Sur. Rep. (Thailand)*, 1989, **20**, 376–389.
13. Sirisopena, N., Torugsa, K., Carr, J., Jugsudee, A., *et al.*, Prevalence of HIV-1 infection in young men entering the Royal Thai Army, Paper presented at the Ninth International Conference on AIDS, Berlin, June 1993 (abstract PO-C08-2778).
14. Sittitrai, W., Phanuphak, P., Barry, J. and Brown, T., Thai sexual behavior and risk of HIV infection: A report of the 1990 survey of partner relations and risk of HIV infection in Thailand, Research Report, Thai Red Cross Society, Programme on AIDS, Bangkok.
15. Mastro, T. D., Satten, G. A., Nopkesorn, T., Sangkharomya, S. and Longini Jr., I. M., Probability of female-to-male transmission of HIV-1 in Thailand, *Lancet*, 1994, **343**, 204–207.
16. NESDB (National Economic and Social Development Board) Working Group, *Projections for HIV in Thailand 1987–2005 An Application of EPIMODEL*, Bangkok, 1994.
17. Duerr, A., Xia, Z., Nagachinta, T., Tovanabutra, S., Tansuhaj, A. and Nelson, K., Probability of male-to-female HIV transmission among married couples in Chiangmai, Thailand, Paper presented at the Tenth International Conference on AIDS and STDs, Yokohama, Japan, August 1994 (abstract 105C).
18. Cash, R.A., Heterosexual behavior related to the risk of HIV infection among Japanese men in Bangkok, Thailand, Paper presented at the Xth International Conference on AIDS and STDs, Yokohama, Japan, August 1994 (abstract PD0367).
19. Choi, K.-H., Catania, J., Coates, T. J., Hyung, L. and Hearst, N., International travel and AIDS risk in South Korea. *AIDS*, 1992, **6**, 1555–1557.
20. Ruan, F. F., *Sex In China: Studies in Sexology in Chinese Culture*, Plenum Press, New York, 1991.

21. Boonchalaksi, W. and Guest, P., *Prostitution in Thailand*, Institute for Population and Social Research Publication No. 171, Bangkok, Thailand, Mahidol University, 1994.
22. Mane, P. and Maitra, S. A., *AIDS Prevention: The Sociocultural Context in India*, Tata Institute of Social Sciences, Bombay.
23. Mawar, N., Divekar, A. D., Tripathy, S. P., Bagul, R. G., Swamy, M., Banerjee, K. and Rodrigues, J. J., Commercial sex workers' risk to HIV in Pune, Paper presented at the Tenth International Conference on AIDS and STDs, Yokohama, Japan, August 1994 (abstract 448D).
24. Stephens, E. J., *et al.*, Intervention development in areas of traditional prostitution in Tamil Nadu, Paper presented at the Tenth International Conference on AIDS and STDs, Yokohama, Japan, August 1994 (abstract PD0447).
25. Carballo, M., Cleland, J., Carael, M. and Albrecht, G., A cross national study of patterns of sexual behavior. *J. Sex Res.* 1989, **26**, 287–299.
26. Sundhagul, D. and Gill, Z. J., Phnom Penh AIDS Awareness Project, World Vision International Cambodia and Office of Population Technical Assistance, Phnom Penh.
27. Seghal, P. N. and Singh, K., *Knowledge, Attitudes, Beliefs and Practices (KABP): Study Related to AIDS in Manipur State, India and Intervention Strategies*. Voluntary Health Association of India, New Delhi, 1992.
28. Lui, P. K., Kwong, P. Lau, J., Cheng, H., Yeung, V. and Wong, T., *The Knowledge, Attitude, Behaviour and Practice Survey on AIDS in Hong Kong 1992*. Hong Kong AIDS Foundation, Hong Kong, 1993.
29. Task Force on the Study of Adolescent Sexuality, 1989, *Working Report on Adolescent Sexuality Study*, 1986, Hong Kong, Family Planning Association of Hong Kong, 1989.
30. Munakata, T. and Tajima, K. The Japanese risk behaviours and their HIV/AIDS preventive behaviours, unpublished report. (Some results from this study are also reported in Munakata, T., *AIDS in Japan*, Akashi Shoten, Tokyo, 1994.
31. Htoon, M.T., Lwin, H. H., San, K. O., Zan, E. and Thwe, M., HIV/AIDS in Myanmar, *AIDS*, 1994, **8** (suppl. 2), S105–S109.
32. Gurubacharya, V. L. and Suvedi, B. K., Sexual behaviour pattern in Nepal, Paper presented at the Tenth International Conference on AIDS and STDs, Yokohama, Japan, August 1994 (abstract PC0123).
33. Suvedi, B. K., Gurubacharya, V. L. and Thapa, K., Seasonal migration and its relation to HIV transmission in western Nepal, Paper presented at the Tenth International Conference on AIDS and STDs, Yokohama, Japan, August 1994, (abstract 192D).
34. Iwamuro, S., Use of condoms among Japanese high school students, Paper presented at the Tenth International Conference on AIDS, Yokohama, Japan, August 1994 (abstract PD0590).
35. Lau, J., *Working Report Evaluation of Educational Programmes on AIDS in Secondary Schools in Hong Kong*, Hong Kong, Advisory Council on AIDS, 1994.
36. Havanon, N., Bennett, A. and Knodel, J., Sexual networking in provincial Thailand, *Studies in Family Planning*, 1993, **24**, 1–17.
37. Xenos, P., Pitaktepsombati, P. and Sittitrai, W., Partner patterns in the sexual behaviour of unmarried, rural Thai men, *Asian Pacific Population Forum*, 1993, **6**(4), 104–117.
38. Hanenberg, R., Rojanapithayakorn, W., Kunasol, P. and Sokal, D., The impact of Thailand's HIV-control programme as indicated by the decline of sexually transmitted diseases. *Lancet*, **344**, 243–245.
39. Merati, T. P., Suarmiartha, E., Ruddick, A. C., Ekstrand, M. and Mandel, J., Assessing the role of traditional Balinese youth groups (STTs) in AIDS prevention programmes, Paper presented at the Tenth International Conference on AIDS, Yokohama, Japan, August 1994 (abstract PD0351).

The economics of the AIDS epidemic: Lessons from the United States

Sherry Glied

Columbia School of Public Health and Department of Economics, Columbia University, 600 West, 168 Street, New York, NY 10027, USA

THE AIDS epidemic officially began in 1981, when a number of physicians in New York and California described the existence of a terrible new disease that was afflicting some of their homosexual patients[1,2]. Later epidemiological research documented cases of infection by the human immunodeficiency virus that had occurred in the United States before 1981 and in countries other than the United States[3]. Today, HIV infection has spread around the globe. Nonetheless, in the thirteen years since those New York and California cases, more AIDS cases have been diagnosed and officially reported in the United States than in any other country. The AIDS cases reported in the United States through 1992 accounted for 44% of all documented AIDS cases for that year[4]. Estimates of the number of AIDS cases in Africa far exceed the US figures, but many of these cases are not reported[3].

Together with the large number of cases diagnosed, the research funding capacity of the United States government and the encouragement of vocal groups of people with AIDS have made the United States the predominant setting for AIDS-related research. In 1990 alone, the United States federal government spent more than $1.6 billion dollars for AIDS-related research and control activities, including substantial funding for social and behavioural research[3]. More than 500 articles have been published on economic, historical and ethnographic aspects of the epidemic in the United States.

The course of HIV disease varies little, whether a case is diagnosed in the United States or in a developing country[5]. As a consequence, medical researchers around the world have been able to benefit from US investments in AIDS research. More than most other diseases, AIDS raises complex issues that cut across economics, sociology, organizational behaviour, and psychology as well as medicine. The nature of HIV transmission, the long incubation period of the virus, and the devastating impact of the illness strongly suggest that economic and other social science research is critical to understanding and ultimately conquering this disease. While at first glance the disease's economic and social implications appear to be quite different in developing and industrial countries, some of the results of the many social science research projects in the United States are relevant for Asia too. This paper describes twelve lessons learnt from such research.

Estimates of the cost of treating HIV infection should incorporate both AIDS and pre-AIDS costs

Before 1993, the clinical definition of AIDS in the United States included only those with specific symptoms of the disease. In 1993 the surveillance definition was broadened to include those with severely depressed immune systems whether or not they manifested particular symptoms[6]. Under the new Centers for Disease Control definition of AIDS, patients' HIV status and CD4 counts (a measure of immune status) alone may determine a diagnosis of AIDS. This adjustment in the surveillance definition has had a substantial impact on the number of cases reported. In New York City, for example, researchers expect a 50% increase in cases among injecting drug users because of the change in definition[7]. Outside the United States, the costs of testing for serostatus and CD4 counts may limit the effect of any broadening of the case definition, leaving many HIV-infected people with severe immunosuppression outside the official count of AIDS cases.

The US experience suggests that estimates of the cost of treating HIV infection that use earlier definitions and begin counting costs only after AIDS symptoms begin may significantly underestimate the full medical costs of the disease. Based on the pre-1993 definition, an estimated 30% of the total HIV-related costs in the United States were for HIV-infected patients who did not have symptoms of AIDS[8]. US estimates suggest that the annual costs of treating HIV infection prior to the development of AIDS symptoms are between 10 and 25% as high as the cost of treating AIDS[9, 10] and that 40% of the lifetime per person cost of HIV infection is incurred before the development of AIDS[11]. Estimates of the proportion of total costs incurred in treating presymptomatic HIV infection in the United States may exceed comparable estimates for other countries because a portion of the US cost of treating presymptomatic patients covers costly prophylactic drug therapies. Estimates that predate the widespread use of such therapies, however, also describe significant costs incurred prior to AIDS diagnosis for patients with wasting syndrome, lymphadenopathy, and other pre-AIDS consequences of HIV infection[12,13].

These studies suggest that financial estimates of the effect of HIV on the health sector should move beyond measuring the costs of AIDS alone. Especially in areas where HIV testing is not done routinely, health care costs may rise substantially before large numbers of AIDS cases have been diagnosed and before the spread of the epidemic is recognized.

Estimates of the cost of the HIV epidemic should incorporate estimates of the cost of related opportunistic infections

The HIV epidemic in the United States has been associated with a resurgence of other infections, particularly tuberculosis[14, 15]. Tuberculosis, an airborne disease, is far more infectious than HIV[3]. HIV-infected patients, with their depressed immune systems, are especially vulnerable to tuberculosis. As the number of HIV-infected people with tuberculosis increases, others, even those who are not HIV-infected, run a greater risk of encountering tuberculosis carriers and becoming infected.

Tuberculosis is a costly disease. For those with non-drug-resistant strains, the estimated cost (inflated using the medical care component of the US consumer price index) of hospital care in the United States averaged $9000 (in 1991 dollars) in one 1980 study[16]. Higher rates of infection, especially in populations that may not complete the course of drug treatment, raise the spectre of drug-resistant tuberculosis. Drug-resistant strains are far more difficult and costly to treat successfully. In the United States the cost of treating drug-resistant tuberculosis can reach as much as $200,000 per case[17].

Tuberculosis was almost eradicated in the United States before the beginning of the HIV epidemic[17]. The dramatic recent growth in cases, especially in centres of HIV infection, raises the possibility that the HIV epidemic may ignite additional subepidemics of costly infectious diseases. If HIV sets back progress in the control of other infectious diseases, the total costs associated with the virus may dwarf those that economists have computed until now.

Health planning, particularly the development of alternative care facilities, can substantially reduce the cost of AIDS care

The earliest estimates of the costs of HIV infection in the United States documented an enormous disparity between costs in the two major centres of infection, New York and San Francisco[18]. Costs of treatment in New York were almost twice as high as costs in San Francisco. This difference has persisted over time and cannot be attributed to differences in the characteristics of AIDS patients in the two cities. Rather, the variation stems from the alternative treatment patterns employed in San Francisco and New York[19].

The less costly pattern of treatment in San Francisco arose because local homosexual men's groups rapidly organized outpatient treatment and hospice facilities for people with AIDS[20]. By contrast, few such community-based out-of-hospital facilities were available in New York[21]. The presence of community-based facilities in San Francisco enabled doctors and hospitals to establish patterns of care that relied on outpatient and nonhospital therapy[22–24]. Despite efforts to establish alternative forms of care for people with AIDS in New York, patients in New York still spend substantially longer time in hospital than do those in San Francisco and are much more likely to die in hospital[19, 25].

Even within cities, patients who have access to out-of-hospital facilities incur lower costs than those who do not. Within New York, the costs of care for injecting drug users without stable housing are significantly greater than those for similar patients with stable housing facilities[26]. The role of social support in reducing costs is particularly evident in the case of paediatric HIV infection. Mothers infected with HIV may not be able to care for their HIV-infected babies and may leave them in the hospital. The length of stay in the hospital for these babies may be as much as four times as long as for babies born with HIV infection who have homes[27].

In response to the growing number of HIV-infected babies in US hospitals, health planners and community groups have developed special foster care facilities[28, 29] for HIV-infected babies. Like the alternative treatment and care facilities developed for adults in San Francisco, these foster care facilities for babies have the potential of reducing the cost of care for HIV patients without diminishing the quality of care provided. These innovations highlight the importance of health planning and interorganizational coordination in developing cost-effective forms of care for people with AIDS.

The cost of AIDS control should be incorporated into estimates of the economic effect of the epidemic

The costs of treating AIDS are so high that investments in prevention and control will frequently be cost-beneficial. The costs of these investments, however, are real resource costs associated with the epidemic. In the United States, these costs include the cost of AIDS research, which in 1991 came to $805 million at the National Institutes of Health alone[30], the costs of tracking the disease, which in 1991 came to $509 million for the Centers for Disease Control[3], the costs of educational campaigns, which in 1991 totalled $700 million for federal and state educational programmes[31], and the cost of precautions to avoid HIV infection. In

health care settings, the cost of implementing the precautions recommended by the federal government amount to more than $800 million per year[31].

These budget items by no means exhaust the range of national efforts for AIDS prevention and control, but these alone, taken together, cost about as much as the lifetime cost of treating 23,000 HIV-infected people[11]. Exclusion of the costs of prevention and control from estimates of the cost of the epidemic leads to significant understatement of its economic impact.

Concentrating AIDS care in a few centres may improve quality and reduce cost

A series of recent studies in the United States have examined the relationship between hospitals' experience in treating HIV patients and the quality of outcomes achieved. In keeping with results from studies of other diseases, researchers have found that hospitals with more experience in treating AIDS generally achieve better results[32]. Mortality rates for patients with a variety of AIDS-related opportunistic infections are significantly lower in hospitals that have treated many AIDS patients[33, 34]. While these studies do not measure the cost of care, other studies have examined changes in the cost of care over the course of the epidemic. These studies suggest that as health providers and health systems learn about HIV the cost of treatment declines. In particular, over a period of time, providers are more likely to have access to and to treat HIV using outpatient treatment facilities[11].

The need for special precautions to lessen the risk of HIV transmission to health care workers also supports a policy of concentrating care in a few centres. Precautions against transmission may involve the use of special equipment and procedures. The fixed costs of acquiring new equipment and adopting new procedures in a hospital may be considerable[35]. Concentrating cases can reduce the average cost of precautions per case. Concentrating cases may also enhance the value of these precautions by encouraging health care workers to cooperate with safety guidelines. Some evidence suggests that compliance with safety precautions is greatest in areas with the highest prevalence of AIDS cases – where medical workers are at highest risk[36].

Concentrating facilities for care of HIV patients in a few centres may improve quality and reduce costs. As physicians gain experience with the disease, they may choose fewer, more effective treatments, they may have a better sense of when to discharge a patient from hospital, and they may know of more alternative facilities that can take patients. Gains from experience across hospitals, though, do not seem to translate into gains from concentrating patients into dedicated units within hospitals. Experiments with dedicated AIDS units show little improvement in costs or quality of care relative to patients cared for in other parts of the same hospital[37].

The possible presence of economies of scale in AIDS treatment suggests that careful health facility planning can reduce the epidemic's costs and suggest a possible source of upward bias in the existing estimates of AIDS costs. Simple linear extrapolations from the cost of treating the first few cases encountered in a country may greatly overstate the full cost of well-organized treatment.

Health manpower planning should incorporate the effects of HIV on provider recruitment

Health care providers who serve patients with HIV infection risk exposure to the virus. Unlike other conditions to which health providers may become exposed, HIV infection cannot be cured or avoided through vaccination. Given the devastating impact of HIV transmission, health providers' reluctance to treat patients with the disease is not surprising.

Problems in personnel recruitment have increased over time. Early in the epidemic, awareness of the risk of transmission was limited[38]. Furthermore, medical personnel may have been interested in being on the forefront of treatment of a new disease. Since then, despite costly precautions to avoid HIV transmission, some medical workers have been infected through occupational exposure[39]. As the number of cases and perceptions of the risk of transmission have increased and the lure of medical innovation has diminished, medical personnel recruitment has been affected[38].

In recent polls, many doctors express an aversion to treating AIDS patients and state that they would not care for these patients if they had a choice[40, 41]. Hospitals with large AIDS caseloads have had trouble recruiting medical staff[42]. Similarly, interest in specialties that now involve caring for many AIDS patients appears to have declined[43].

The cost implications of recruitment difficulties can be monetized as the compensating differential need to attract providers to care for HIV patients. Studies have not yet estimated the magnitude of this compensating differential, but polling data and recruitment evidence suggest that it may be substantial. Health planners should take into account the possibility that recruitment of personnel will become increasingly difficult and costly as the epidemic progresses.

Prevention efforts should be targeted to high-risk groups and areas

The first cases of AIDS reported in the United States were among homosexual men and intravenous drug users. Thirteen years later, these two groups continue to

account for about 80% of all newly diagnosed cases[44]. The epicentres of the epidemic in the early 1980s were San Francisco, Los Angeles and New York City. Today, these three cities account for almost a quarter of newly diagnosed cases. Although the epidemic has spread beyond the original risk groups and to other regions of the country, it remains heavily concentrated in the populations and places where it began. The growing number of cases among heterosexuals can often be traced back to contact with a member of the traditional risk groups. This epidemiological pattern suggests that the benefits of prevention and control efforts targeted at traditional risk groups may persist over time.

A growing body of evidence also suggests that prevention and control strategies targeted at lower risk groups may be quite ineffective. Despite considerable concern over the possible spread of the AIDS epidemic in the United States to groups other than homosexual men and injecting drug users, few changes in the behaviour of low-risk groups have been documented. While behaviour changes rapidly in response to perceived increases in risk in local communities, perceptions of general risk do not seem to affect behaviour[45]. Changes in behaviour have been greatest among urban homosexual men, followed by injecting drug users. Lower-risk groups, particularly young heterosexuals, although they are quite well-informed about HIV, show few changes in behaviour in response to the epidemic[46]. Costly efforts to encourage changes in behaviour in these groups, including public educational campaigns, have had limited effects on behaviour. While some increase in condom use among heterosexuals has occurred, most of this change appears to have followed early media reports about the epidemic. US Surgeon General Koop's public mailing of AIDS information to all American households in 1988 did not affect measured risk behaviours significantly[47].

Most educational campaigns for AIDS prevention among heterosexuals in the United States encourage condom use. Condoms are inexpensive relative to income in the United States, and in some schools and universities are provided free. Nonetheless, as recently as 1988, fewer than 20% of never-married women in the United States who reported using contraception used condoms[48].

These results suggest that the nonmonetary costs of changing sexual behaviour are high. Unless people perceive a significant risk of infection, they may rationally choose not to change their sexual practices[31]. Combined with the evidence that the epidemic has remained concentrated at its original sites, the reluctance of those at low risk to change their behaviour suggests that the most cost-effective approach may be to target prevention and control efforts at high-risk groups.

Groups that face a high risk of infection will change their behaviour if they have information and opportunity

In 1986 researchers forecast that 87,000 new AIDS cases would be diagnosed in the United States in 1991. The actual number of cases diagnosed in 1991 missed this mark by 41,000. The reduction in AIDS incidence by 1991 occurred because groups at high risk of HIV infection had made massive changes in their behaviour to avoid infection[49].

Estimates of behaviour change suggest that news about AIDS led to a rapid and profound change in the sexual behaviour of homosexual men[50]. One report finds that the proliferation of information about the epidemic resulted in a threefold to fourfold decrease in rates of high-risk sex among this group[51]. Studies of injecting drug users also find evidence of substantial changes in behaviour. Towards the end of 1983, drug users had reduced needle-sharing and increased their efforts to clean needles between injections[52].

These results suggest that even before major government campaigns, individuals who faced a significant risk of infection took steps to reduce their vulnerability. For homosexual men, these changes in behaviour led to a notable reduction in the spread of the disease. By 1991, cases among homosexual men were about one-third lower than had been projected[49]. Unfortunately, behavioural change among injecting drug users had smaller effects on the spread of the epidemic among this population. Effective behavioural change for members of this group required access to costly goods and services – drug treatment programmes, bleach, clean needles – that may have been unaffordable given their incomes. Public funding for programmes aimed at injecting drug users has been mired in controversy and slow to respond to the epidemic[53].

Evidence from high-risk groups in the United States suggests that information about HIV can spread more rapidly than the virus itself. People who perceive themselves as vulnerable to infection do make efforts to change their behaviour. For many, however, public support is necessary if these behavioural changes are to be translated into reductions in infections.

Complacency is dangerous, behaviour changes are difficult to maintain

While the evidence of changes in behaviour among high-risk groups is encouraging, some recent studies document increases in dangerous behaviours among these groups. These studies are the mirror image of those that found reductions in high-risk behaviour early in the epidemic. Then, as the perceived risk of infection increased, vulnerable populations changed their behaviour. Now, as the perceived risk of infection

declines, vulnerable populations show signs of resuming their preepidemic behavioural patterns[31 50].

High rates of risky sexual behaviour have been documented among homosexual men in small cities, where AIDS rates are low, and among young homosexual men[54]. The HIV epidemic has not led to a fundamental change in sexual habits. Rather, homosexual men in communities that have not been severely affected by the epidemic, whether because they are geographically isolated or because they include men who have come of age since 1981, continue to practice high-risk sex. Some studies also document reversions to high-risk behaviour among homosexual men who had changed their behaviour earlier in the course of the epidemic. These studies are consistent with the view that the nonmonetary cost of safe sex is high, and that only those who perceive a significant risk of developing AIDS will find that the benefits of changing behaviour outweigh the costs. Any reductions in high-risk behaviour should be cheered, but health planners should beware of the potential for recidivism.

Community participation should be encouraged

Among both homosexual men and injecting drug users, significant information dissemination leading to behaviour change occurred through community action. Community organizations sponsor educational activities and take an active role in marshalling public funding for AIDS research and prevention[20]. The epidemic's focus on traditionally marginalized groups, who may not trust public authorities, makes the role of community groups in education especially important.

Community organizations have also been at the forefront of treatment innovations. The development of out-of-hospital facilities for the care of AIDS patients was largely the work of voluntary community groups. Especially in the first decade of the epidemic, volunteers provided a significant portion of the care patients received[20]. Community organizations and voluntary labour have substantially reduced the cost of preventing AIDS and caring for AIDS patients in the United States.

Some researchers have expressed concern that the potential of community organizations and volunteers may soon be exhausted[20]. These organizations draw staff and resources from the very communities that have been hardest hit by the epidemic. As more community members succumb, the burden on those who remain healthy may become excessive. Without the participation of community organizations, the quality of educational interventions and of care offered to those with AIDS will suffer, and the monetary costs of these activities will rise.

The US experience suggests that community groups can contribute greatly to the task of coping with the HIV epidemic, but that these groups can become exhausted. Health planners should encourage the development of community groups and their participation in AIDS prevention and care. They should also provide both encouragement and tangible support to these groups as the epidemic continues.

Efforts to reduce stigma and discrimination against people with HIV can reduce the cost of the epidemic

HIV is not infectious in ordinary workplace or educational settings. Unfortunately, fears and misconceptions about the disease abound. In 1987 one-third of Americans polled believed that an HIV-infected worker should be fired and one-quarter stated that they would not work with a person who had AIDS[3]. Parents throughout the United States have tried to keep HIV-infected children out of schools[5]. Such attitudes may lead to unwarranted and costly dismissals of productive people and to increased costs for child care and education. On average, HIV-infected people carry the virus for ten years before manifesting symptoms. The present value of foregone income over the remaining healthy lifetime of an average American male full-time worker who loses his job when he becomes HIV-infected is almost $250,000 (based on mean money income in 1991 of year-round, full-time male worker discounted at 5% over ten years [55]. Estimate of healthy life expectancy is reported in ref. 3.).

Efforts to educate the general public about the realities of HIV may have succeeded in reducing the stigma associated with AIDS[56]. By 1991, the number of people reporting that they believed an HIV-infected worker should be fired and that they would not work with a person who had AIDS had dropped to one in five and one in six, respectively[3]. Nonetheless, considerable stigma and misinformation remain[57]. Information campaigns designed to reduce discrimination may pay off in better use of productive resources.

Economic policy should redistribute the burden of HIV costs from affected regions to the nation as a whole

The lifetime cost of treating HIV infection in the United States is more than $100,000 (ref. 11). The almost 145,000 people who contracted AIDS in the United States during the 1980s were predominantly young, productive men with many years of earnings ahead of them[44]. Nor are there readily available substitutes for these men. During the last half of the 1980s, unemployment rates in the United States for men over 20 averaged just over 5% (ref. 56). Despite the high cost of treatment, the productivity of those who contracted

AIDS and the relatively inelastic supply of labour in the United States, the epidemic has not had a devastating effect on the macroeconomy of the nation as a whole. The estimated lifetime federal and state tax revenue losses associated with the 35,000 Americans who died in 1992, for example, amounted to only 1.5% of the revenue of the federal government in that year[1, 5, 44]. The medical care costs of AIDS account for a similar portion of total United States health expenditures[13].

The relatively benign effect of the epidemic on the macroeconomy of the nation as a whole contrasts with its devastating effects on particular regions, especially on communities within these regions[18, 20]. In New York City and San Francisco, the total number of AIDS cases diagnosed since the beginning of the epidemic amount to more than 1 case for every 150 current residents[44]. In some New York City neighbourhoods, the number of AIDS cases diagnosed in the 1980s amounted to one for every 55 residents[20]. The costs of treatment have wrecked the finances of local hospitals and forced state policy-makers to make difficult decisions about resource allocation[20, 58].

The nature of transmission of HIV suggests that the epidemic will remain concentrated in particular areas of the country and in particular communities. The gulf between the ruinous effects of the epidemic on these areas and communities and its less severe effects on the United States macroeconomy highlights the importance of redistributive policies in coping with the epidemic. Effective policies will help these communities bear the epidemic's social and medical costs and, through prevention and control measures, reduce the spread of the epidemic to other areas. Thus, both compassion and self-interest dictate that nations assist those areas suffering under the weight of the epidemic. At least in the United States the economic effects of the epidemic spread across the nation as a whole will easily permit such redistributive actions.

Conclusions

Economic and social research in the United States can improve our understanding of the epidemic in other parts of the world. In terms of measuring the costs of the epidemic, United States research suggests that conventional methods of accounting may overestimate some aspects of costs by failing to recognize cost-saving methods of care, including outpatient therapy and treatment centres, that may develop as the epidemic progresses. It also suggests that many measures of economic impact often seriously underestimate the costs of the epidemic by failing to include pre-AIDS costs, the costs of related opportunistic infections, the costs of HIV prevention and control, and the increased cost of recruiting medical personnel. American research provides cause for both hope and alarm in terms of HIV prevention. High-risk groups, particularly those in active and well-supported communities, have reduced their levels of dangerous behaviours remarkably quickly. Furthermore, educational efforts may be effective in reducing the level of discrimination people with AIDS encounter. By contrast, reversions to risky behaviour are common even in groups that had succeeded in changing their behaviour, and groups that perceive themselves to be at low risk have hardly responded to the threat of AIDS. Perhaps the most important lesson of the United States experience is that the effects of the disease are highly localized. National policy must be targeted at discrete, severely affected communities.

Lessons from the United States may be valuable. At the same time, one must recognize that the United States epidemic differs substantially from that in other countries. Most people with AIDS in the United States are homosexual men and drug users. Prostitution and heterosexual sex are not important vectors of transmission in the United States. The structure of economies and labour markets varies profoundly among the three major regions of HIV prevalence: the United States, Sub-Saharan Africa, and Asia. Cultural factors, including the role of community organizations and social structure, differ immensely by country, and these play a critical part in determining how quickly the spread of the epidemic can be arrested and how much it will cost to care for those who become ill. These differences may diminish the transferability of United States knowledge on the economic and social effects of AIDS, but as the epidemic rages, it makes sense for researchers throughout the world to glean what they can from each other's work.

1. Centers for Disease Control, Kaposi's sarcoma and Pneumocystis pneumonia among homosexual men – New York City and California, *Morb. Mort. Weekly Rep.*, 1981, **30**, 305–308.
2. Gottlieb., M. S., Schroff, R., Schanker, H. M., Weisman, J. D., Phan, P. T., Wolf, R. A. and Saxon, A. S., *Pneumocystis carnii* pneumonia and mucosal candidiasis in previously healthy homosexual men: Evidence for a new acquired cellular immunodeficiency, *New Engl. J. Med.*, 1981, **305**, 1425–1431.
3. Hombs, Mary Ellen. *AIDS Crisis in America*, ABC-CLIO, *Santa Barbara*, 1992.
4. WHO *Weekly Epidemiological Record,* July 1992.
5. Sills, Yole G., *The AIDS Pandemic*, Greenwood Press, Westport, 1994.
6. Centers for Disease Control and Prevention, HIV/AIDS surveillance report, 1993.
7. Des Jarlais, D. C., Wenston, J., Friedman, S. R., Sotheran, J. L., Maslansky, R., Marmor, M., Yancovitz, S. and Beatrice. S., Implications of the revised surveillance definition: AIDS among New York City drug users, *Am. J. Publ. Health*, 1992, **82**, 1531–1533.
8. Andrulis, D. P., Weslowski, V. B., Hintz, E. and Spolarich, A. W., Comparisons of hospital care for patients with AIDS and other HIV-related conditions, *JAMA*, 1992, **267**, 2482–2486.

9. Hay, J. W., Osmond, D. H. and Jacobson, M.A., Projecting the medical costs of AIDS and ARC in the United States, *J. AIDS*, 1988, **1**, 466–485
10. Hellinger, F. J., Forecasts of the costs of medical care for persons with HIV: 1992–1995. *Inquiry* Fall, 1992, **29**, 356–365.
11. Hellinger, F. J., The lifetime cost of treating a person with HIV, *JAMA*, 1993, **270**, 474–478.
12. Scitovsky, A. A., Past lessons and future directions: The economics of health services delivery for HIV-related illnesses. Presented at NCHSR conference May 17–19 , 1989 and Atlantic Information Services Inc., *AIDS Information Guide*, February 1990, #313.
13. Scitovsky, A. A., Studying the cost of HIV-related illnesses: reflections on the moving target, *Milbank Quarterly*, 1989, **67**, 318–344.
14. Braun, M. M., Cote, T. R. and Rabkin, C. S., Trends in death with tuberculosis during the AIDS era, *JAMA*, 1993, **269**, 2865–2868.
15. Frieden, T. R., Sterling, T., Pablos-Mendez, A., Kilburn, J. O., Cauthen, G. M. and Dooley, S. W., The emergence of drug-resistant tuberculosis in New York City, *New Engl. J. Med.*, 1993, **328**, 521–526
16. Rose, R.C., and McGowan Jr., J. E., Urban tuberculosis: some modern problems, *Am. J. Med. Sci.*, 1984, **287**, 24–26.
17. Perutz, M. F., The White Plague, *The New York Review of Books*, 1994 XVI, No. 10, pp. 35–39.
18. Bloom, D. E. and Carliner, G., The economic impact of AIDS in the United States, *Science*, 1988, **239**, 604–610.
19. Andrews, R. M., Keyes, M. A., Fanning, T. R. and Kizer, K. W., Lifetime Medicaid service utilization and expenditures for AIDS in New York and California, *J. AIDS*, 1991, **4**, 1046–1058.
20. *The Social Impact of AIDS in the United States.* (eds Jonsen, A. R. and Stryker, J.), National Academy Press, Washington, DC 1993, p. 160.
21. Shulman, L. C. and Mantell, J. E., The AIDS crisis: a United States health care perspective, *Social Science and Medicine*, 1988, **26**, 979–988.
22. Beresford, L., Alternative, outpatient settings of care for people with AIDS, *Quality Review Bull.*, 1989, **15**, 9–16
23. Martin, J. P., Hospice and home care for persons with AIDS/ARC: meeting the challenges and ensuring quality. *Death Studies*, 1988, **12**, 463–480.
24. Volberding, P. A., Caring for the patient with AIDS. An integrated approach, *Infectious Disease Clinics of North America*, 1988, **2**, 543–550.
25. Kelly, J. J., Chu, S. Y. and Buehler, J. W., AIDS deaths shift from hospital to home. AIDS mortality project group, *Am. J. Publ. Health*, 1993, **83**, 1433–1437.
26. Bennett, C. L., Pascal, A., Cvitanic, M., Graham, V., Kitchens, A. and DeHovitz, J. A., Medical care costs of intravenous drug users with AIDS in Brooklyn, *J. AIDS*, 1992, **5**, 1–6.
27. Hegarty, J. D., Abrams, E. J., Hutchinson, V. E., Nicholas, S. W., Suarez, M. S. and Heagarty, M. C., The medical care cost of human immunodeficiency virus-infected children in Harlem, *JAMA*, 1988, **260**, 1901–1905.
28. Laxton, C. E., A foster home for HIV infants. Grandma's House, *Caring*, 1989, **8**, 37–39.
29. Rendon, M., Gurdin, P., Bass, J. and Weston. M., Foster care for children with AIDS: a psychosocial perspective, *Child Psychiatry & Human Development*, 1989, **19**, 256–269.
30. Insitute of Medicine, *The AIDS Research Program of the National Institutes of Health*, National Academy Press, Washington DC, 1991.
31. Philipson, T. J. and Posner, R. A., *Private Choices and Public Health.* Harvard University Press, Cambridge, MA, 1993.
32. Sandrick, K., Learning from experience, In AIDS treatment, knowledge means quality, *Hospitals*, 1993, **67**, 32–35.
33. Bennett, C. L., Garfinkle, J. B., Greenfield, S., Draper, D., Rogers, W., Mathews, C. and Kanouse, D. E., The relation between hospital experience and in-hospital mortality for patients with AIDS-related PCP, *JAMA*, 1989, **261**, 2975–2979
34. Turner, B. J. and Ball, J. K., Variations in inpatient mortality for AIDS in a national sample of hospitals, *J. AIDS*, 1992, **5**, 978–987.
35. Stock, S. R., Gafni, A. and Bloch, R. F., Universal precautions to prevent HIV transmission to health care workers: an economic analysis *Can. Med. Assoc. J.* 1990, **142**, 937–946.
36. Kaczmarek, R. G., Moore Jr., R. M., McCrohan, J., Arrowsmith-Lowe, J. T., Caquelin, C., Reynolds, C. and Israel, E., Glove use by health care workers: results of a tristate investigation, *Am. J. Infection Control*, 1991, **19**, 228–232.
37. Fahs, M. C., Fulop, G., Strain, J., Sacks, H. S., Muller, C., Cleary, P. D., Schmeidler, J. and Turner, B., The inpatient AIDS unit: a preliminary empirical investigation of access, economic, and outcome issues. *Am. J. Publ. Health*, 1992, **82**, 576–578.
38. Brennan, T. A., Occupational transmission of HIV: an ethical and legal challenge, in *AIDS and the Health Care System* (ed. Gostin, L. O.), Yale University Press, New Haven, 1990, 143–154.
39. Bell, D. M., HIV infection in health care workers: Occupational risk and prevention, in *AIDS and the Health Care System.* (ed. Gostin, L. O.), Yale University Press, New Haven, 1990, 115–125, p. 118.
40. Hayward, R. A. and Weissfeld, J. L., Coming to terms with the era of AIDS: attitudes of physicians in US residency programs, *J. Gen. Int. Med.*, 1993, **8**, 10–18
41. Shapiro, M. F., Hayward, R. A., Guillemot, D. and Jayle, D., Residents' experiences in, and attitudes toward, the care of persons with AIDS in Canada, France, and the United States, *JAMA*, 1992, **268**, 510–515.
42. Ness, R. B., Kelly, J. V. and Killian, C. D., House staff recruitment to municipal and voluntary New York City residency programs during the AIDS epidemic, *JAMA*, 1991, **266**, 2843–2846.
43. Ness, R. B., Killian, C. D., Ness, D. E., Frost, J. B. and McMahon, D., Likelihood of contact with AIDS patients as a factor in medical students' residency selections, *Academic Medicine.* 1989, **64**, 588–594.
44. Centers for Disease Control and Prevention, HIV/AIDS Surveillance Report, 1993 revised classification system for HIV infection and expanded surveillance case definition for AIDS among adolescents and adults, JAMA, 1993, **269**, 729–730.
45. Klepinger, D. H., Billy, J. O., Tanfer, K. and Grady. W. R., Perceptions of AIDS risk and severity and their association with risk-related behaviour among US men, *Family Planning Perspectives*,1993, **25**, 74–82.
46. Freudenberg, N., AIDS prevention in the United States: lessons from the first decade, *Int. J. Health Services*, 1990, **20**, 589–599.
47. Singer, E., Rogers, T. F. and Glassman, M. B., Public opinion about AIDS before and after the 1988 US government public information campaign, *Public Opinion Quarterly*, 1991, **55**, 161–179.
48. Mosher, W. D. and Pratt, W. F., *Contraceptive Use in the United States, 1973–1988. Advance Data Number 182,* National Center for Health Statistics, 1990.
49. Bloom, D. E. and Glied, S., Projecting the number of new AIDS cases in the United States, *Int. J. Forecasting*, 1992, **8**, 339–365.
50. Becker, M. H. and Joseph, J. G., AIDS and behavioural change to reduce risk: a review, *Am. J. Publ. Health*, 1988, **78**, 394–410.
51. Fan, D. P., Quantitative estimates for the effects of AIDS public education on HIV infections, *Int. J. Biomed. Computing*, 1993, **33**, 157–177.

52. Des Jarlais, D. C. and Friedman, S. R., HIV infection among intravenous drug users: epidemiology and risk reduction, *AIDS*, 1987, **1**, 67–76.
53. Friedman, S. R., Des Jarlais, D. C. and Ward, T. P., Overview of the history of the HIV epidemic among drug injectors (Ch. 1). in *Handbook on Risk of AIDS: Injection Drug Users and Sexual Partners* (eds Brown, B. S. and Beschner, G. M.), Greenwood Press, Westport, 1993.
54. Kelly, J. A., Murphy, D. A., Roffman, R. A. *et al.*, Acquired imunodefeciency syndrome/human immuno deficiency virus risk behaviour among gay men in small cities, Findings of a 16-city national sample, *Arch. Int. Med.*, 1992, **152**, 2293–2297.
55. Council of Economic Advisers, *Economic Report of the President*. Government Printing Office, Washington, D.C, 1993.
56. Snyder, L. B., The impact of the surgeon general's understanding AIDS pamphlet in Connecticut, *Health Commun.*, 1991, **3**, 37.
57. Herek, G. M. and Capitanio, J. P., Public reactions to AIDS in the United States: A second decade of stigma, *Am. J. Publ. Health*, 1991, **83**, 574–577.
58. Bayer, R., AIDS: the politics of prevention and neglect, *Health Affairs*, 1991, (Spring), 87–97.

Making the blood supply safe from HIV: The Sri Lankan experience

Nandrani S. de Zoysa

14, Bodhiraja Mawatha, Jayanthipura, Battaramulla, Sri Lanka

SRI LANKA is an island situated in the Indian Ocean with an area of 65,610 km^2 and an estimated population of 17.5 million. The National Blood Transfusion Service (NBTS) of Sri Lanka is managed by the government and funded by an allocation from the Ministry of Health budget. This service, which was established in 1962 with a central blood bank and one regional blood bank, has now expanded to include 47 regional blood banks.

The central blood bank, which is located on the premises of Sri Lanka's premier hospital, fulfils the hospital's blood needs, including those of Accident Service and a nearby maternity hospital. It also serves as a reference centre for all the regional blood banks and as a training centre for medical, nursing and technical personnel appointed to the NBTS and their undergraduate and postgraduate trainees. The regional blood banks are located at teaching, provincial, base and district hospitals and handle the blood needs of those hospitals. The central blood bank and regional blood banks also undertake serological investigations, including antenatal serology, for hospitals. The NBTS collects blood from approximately 100,000 donors annually. Facilities for preparing and storing all blood components are available at the central blood bank and eight large regional blood banks, and facilities to prepare a few basic blood components are available in seventeen of the smaller regional blood banks. Blood is tested for syphilis, malaria, hepatitis B and HIV.

Table 1. HIV/AIDS status in Sri Lanka, 31 December 1993

120 HIV/AIDS	→	20 Foreigners
	→	100 Sri Lankans
100 Sri Lankans	→	• 74 Males
	→	* 26 Females
Male:Female ratio		3:1
100 Sri Lankans	→	33 AIDS cases (28 dead)
	→	67 HIV-positive
33 AIDS cases	→	25 Males (22 dead)
	→	8 Females (6 dead)
67 HIV-positives	→	49 Males
	→	18 Females

• 18 Seamen; • 9 Blood donors; * 2 following blood transfusions abroad.

The HIV and AIDS situation

The first case of AIDS in Sri Lanka was detected in September 1986, in a foreigner who became ill while on vacation; the first Sri Lankan with HIV, a male homosexual who became infected when living abroad, was diagnosed in April 1987. Since then, as of 31 December 1993, 120 people had been found to be HIV-positive. Of these, 20 were foreigners and 100 were Sri Lankans (74 men, 26 women). Thirty-three of the Sri Lankans developed AIDS and twenty-eight of them are now dead.

The first HIV-positive blood donor was detected in January 1989, and as of 31 December 1993, a total of nine blood donors had been confirmed as infected with HIV out of a total of some 388,000. All the infected blood donors were men, and most were first-time donors.

Safety of the blood supply

Testing individuals for the HIV antibody cannot assure a safe blood supply, because an individual who donates blood during the period after HIV has entered the body but before the antibody appears in the blood can transmit the virus through blood transfusion. Although 95.0% of HIV-infected people test positive within three months of infection and a further 4.5% test positive within six months, 0.5% remain negative for as long as several years. Therefore, simply testing blood for HIV antibody prior to transfusion will not eliminate the transmission of HIV infection through blood. To make the blood supply as safe as possible, testing for the HIV

Table 2. HIV infection in blood donors in Sri Lanka (20 April 1988 to 31 December 1993)

HIV-antibody-positive donor	Sex	Date of testing	Blood bank where blood was donated	No. of donors tested from April 1988
1	Male	22.01.89	Central Blood Bank (Mobile Unit)	21,973
2	Male	24.04.91	Regional Blood Bank (Kalutara)	137,900
3	Male	03.10.91	Central Blood Bank	173,956
4	Male	12.12.91	Regional Blood Bank (Panadura)	188,681
5	Male	03.08.93	Central Blood Bank	342,848
6	Male	01.09.93	Regional Blood Bank (Castle Street Hospital Colombo)	351,489
7	Male	25.10.93	Regional Blood Bank (Negombo)	367,889
8	Male	05.11.93	Central Blood Bank (Mobile Unit)	372,415
9	Male	27.11.93	Regional Blood Bank (Ragama)	380,000

Table 3. HIV antibody testing in blood banks in the National Blood Transfusion Service 20 April 1988 to 31 December 1993

Year	No. of blood banks testing blood	Percentage of total blood supply covered	No. of blood donors tested	No. of positive blood donors detected
1988	CBB + 13 RBB	25.82	16,426	0
1989	CBB + 18 RBB	60.12	37,686	1
1990	CBB + 32 RBB	91.15	66,353	0
1991	CBB + 36 RBB	99.52	78,543	3
1992	CBB + 41 RBB	99.94	87,461	0
1993	CBB + 43 RBB*	99.96	102,022	5
		Total	388,491	9

*Includes two new RRBs established in 1993.
Prevalence of HIV infection in blood donors = 1 in 43000.

antibody must be considered only as a backup to donor selection, which must be done with meticulous care. The NBTS has adopted a number of strategies to prevent the transmission of the infection to donors during blood donation and to patients through blood transfusion.

HIV infection by blood donation

Instances have been reported from abroad in which blood donors have been infected with HIV during blood donation. This is possible if contaminated lancets, syringes and needles are reused without adequate sterilization. This risk can be eliminated completely by using only disposable sterile items to collect blood. Since 1988 the NBTS has used only disposable sterile items (replacing the earlier practice of reusing items sterilized by high-pressure steam). Therefore, no donor who donates blood at any of the NBTS's 48 blood banks, including mobile units, can contract HIV or any other infection by donating blood. Even the cotton wool and gauze used to clean the blood donor's arm before blood collection is sterilized under high pressure in dressings drums and discarded after use. Sharp bins are used to dispose of lancets, syringes and needles after use to prevent infection through accidental injury to blood donors, blood bank staff and visitors.

HIV infection by blood transfusion

Blood is collected only from voluntary, nonremunerated donors. The former system of partial remuneration was phased out by 1987. However, as in most developing countries, a 'hidden' system exists in which patients asked to get a relative or a friend to donate blood as a replacement for blood that has been or will be transfused to them instead pay money to people to donate

blood for them. This jeopardizes the safety of the blood supply, because these paid donors will not divulge vital information about their health and will donate blood even if they have indulged in high-risk behaviour. The staff responsible for donor recruitment and screening are alerted to this problem during their training and instructed to exclude these donors, or to collect the blood but discard it without informing the donors or the patients if they are unable to reject the donors. Blood for patients is then provided from the stocks available at the blood bank.

During the last five years, the NBTS and Sri Lanka's AIDS Control Programme have arranged confidential circulars, lectures and seminars to instruct staff responsible for blood donor screening and registration about individuals and groups at high-risk for HIV infection.

Blood donors are given a leaflet entitled 'Important message to blood donors – AIDS' which explains about HIV and AIDS infection and appeals for donors to exclude themselves if they belong to one of the high-risk groups listed. Staff are instructed to ask donors if they have read the leaflet prior to registration. If the donor reads the leaflet and goes away, no questions are asked. If the donor has any questions, a confidential interview is arranged with the medical officer. This leaflet was introduced in 1987.

Blood donors have always been required to sign certificates of fitness declaring that they do not have epilepsy, a sexually transmitted disease, tuberculosis, malaria or typhoid. This certificate has just been updated to include heart disease, high blood pressure, stroke, diabetes, cancer, kidney disease, abortion (within six months), and HIV or AIDS. The statements 'I also declare that I have read the leaflet explaining about AIDS and I do not belong to the high-risk groups mentioned in the leaflet' and 'I agree to have my blood tested for HIV (the AIDS virus) and other infections' have been added to this certificate, which the blood donor has to sign prior to registration for blood donation.

NBTS staff have been instructed to examine the forearms of blood donors for injection marks and to reject such donors so as to exclude possible intravenous drug abusers. Staff have also been told to look for skin lesions resembling Kaposi's sarcoma, which is sometimes seen in AIDS patients.

Safer blood and blood products

The NBTS collects blood into citrate phosphate dextrose-adenine blood bags with a shelf life of 35 days, and the blood is processed into blood components. Blood components are cellular and plasma components, which are separated from whole blood by conventional blood bank methods such as centrifuging and freezing. Blood is tested for HIV and other infections prior to transfusion.

NBTS staff question blood donors carefully about their present and previous occupations, travel abroad in the past ten years (including visits to African countries after 1977) and whether they have ever been sailors or in prison. They also look out for foreigners. The earlier practice of rejecting high-risk donors after explaining the problem to them was stopped because blood donors were offended. This was replaced by confidential unit exclusion in late 1991, whereby blood is collected from all donors who are otherwise fit to donate blood, but the blood of high-risk donors is discarded without informing the donor or the patient (if blood was donated for a particular patient), even if the HIV antibody test is negative, and blood for the patient is provided from stocks. Although sometimes a large number of blood units have to be discarded, this has not posed a major problem as yet, because the number of blood donors has increased considerably over the years.

Testing samples of the blood of high-risk donors, i.e. those who tested positive for syphilis or hepatitis B, was started in November 1987, and all 203 samples tested through April 1988 were negative for the HIV antibody.

Routine testing of blood for the HIV antibody using the ELISA technique started in April 1988. By the end of that year 25.8% of the blood collected by the NBTS was tested. This was gradually extended, and currently more than 99.9% of the blood supply is tested for the HIV antibody prior to transfusion. Two regional blood banks are not testing for HIV. These are located in the north in the war zone and cannot obtain a regular supply of test kits because of transport problems. However, these blood banks collect only a few units of blood, which accounts for less than 0.1% of the total blood supply.

At present, blood from the central blood bank and seven regional blood banks is tested for HIV at the anti VD Campaign in Colombo, and two regional blood banks are testing at the anti VD Campaign in Katugastota using the ELISA technique, or the particle agglutination test (Serodia) when ELISA test kits are unavailable. The other regional blood banks perform the HIV antibody test in their own laboratories using either the Serodia test or a quick test (HIV-check or HIV-spot). Trained medical and nursing officers do the HIV antibody test in all but the five regional blood banks that have laboratory technicians. The funds for the HIV antibody test kits were provided initially by the World Health Organization, later by the United Nations Development Programme, and currently are provided by the Asian Development Bank.

If a sample of blood from a donor is found to be HIV-positive, NBTS staff discard the blood even if the blood from the same donor is negative when tested again. If a blood donor is confirmed as being HIV-positive after repeat testing or confirmatory testing using the immunoblot (western blot) technique, this information is

communicated only to the directors of the AIDS Control Programme and the NBTS. The AIDS Control Programme then traces and contacts the HIV-infected blood donor and arranges for counselling. Confidentiality is maintained so that neither the blood bank staff nor the patient (if blood has been donated for a particular patient) are aware that the blood donor has been confirmed as being HIV-infected.

Blood products are plasma fractions that are separated by means of a plasma fractionation unit. This equipment is expensive and is currently not available in Sri Lanka. Therefore, a small quantity of lyophilized factor VIII concentrate (antihaemophilic factor) is imported to treat haemophilia patients who develop reactions to cryoprecipitate prepared locally. To ensure that the antihaemophilic factor is safe, only virus-inactivated products have been imported since 1987. As the products imported from France, Austria and United States before 1987 were not virus-inactivated, a haemophilia register was started in 1988, and haemophilia patients who seek treatment at government hospitals were registered at the central blood bank. As of December 1993, 132 haemophilia patients had been registered, of which 43 had received imported and antihaemophilic factor before 1987. All 132 patients were tested. All tested negative for HIV, but five were found to be positive for hepatitis B.

The NBTS also imports a limited quantity of 5% human serum albumin for patients who develop reactions to plasma transfusions and for use as a replacement fluid for therapeutic plasma exchange. This product is pasteurized and, therefore, is considered safe.

Conclusions

The strategies adopted by the NBTS of Sri Lanka to try to ensure a safe blood supply appear to have been effective because no case of HIV infection following blood transfusion in Sri Lanka has been reported to date. However, with the detection of 100 HIV-infected people in Sri Lanka as of 31 December 1993, and with an estimated 3500 people likely to be infected, resulting in an increased prevalence of HIV infection among blood donors, every effort will have to be made if the NBTS is to continue to provide a safe blood supply.

Ethical issues posed by HIV testing and screening

Ronald Bayer

Columbia School of Public Health, Columbia University, 600 West 168 Street, New York, NY 10027, USA

FROM the outset, the test developed to detect the antibody to the AIDS virus, which was first used on a broad scale in blood banking beginning in 1985, has been mired in controversy[1]. Would groups at increased risk for AIDS be encouraged to take the test? How forceful would such encouragement be? How would those who agreed to be tested be counselled about the test's significance for themselves and others? Would, and could, the results be kept confidential? Would voluntary testing be a prelude to compulsory screening? What would be the consequence of testing for the right to work? To go to school? To bear children? To remain free? Each of these questions would force a confrontation about the relationship between the right to privacy and the protection of the public health and about the roles of voluntarism and coercion in the social response to the threat of AIDS.

Controversy about the HIV test is not foreign to India. The November 1993 issue of the *Indian Journal of Medical Research* published three articles that touched on ethical and policy issues surrounding HIV testing. Mathan[2] argues that the HIV test should be freed of limits imposed by an 'exceptionalist'[3] perspective. 'HIV and AIDS are infectious and should be dealt with exactly as other infectious diseases are being handled. Diagnostic tests are best left to the discretion and judgment of the concerned clinicians[2]. Even this call for a standard of presumed consent, however, does not entail a call for wide-scale mandatory testing. 'Since the availability of HIV testing facilities is limited, the test in any case is not going to be done widely.'

In sharp contrast, Narain *et al.*[4] of the South East Asia Regional Office of the World Health Organization present a strong case for limiting the role of HIV testing in public health practice and for a standard of specific, informed consent because of its unique features: 'Compared to any other type of disease, the issues related to the diagnosis of HIV infection are far more complex.' This plea for 'exceptionalism' concludes by stating: 'Mandatory testing or testing without informed consent is not only counterproductive but also wasteful of scarce resources.'

Finally, Lal and Thakur[5], writing for the National AIDS Control Organization, present proposals then under consideration for its national HIV testing policy document. They too reject testing without 'explicit consent' except for surveillance purposes, where unlinked anonymous screening can assure that no individual is identified.

For an American, writing in a very different ethical, legal, social, economic and epidemiological context to provide simple ethical instruction on HIV testing would be presumptuous. Far more useful, I believe, would be to present a picture of how the ethics of screening have evolved in a variety of settings, thereby alerting readers of *Current Science* to issues that will require attention in the context of a growing AIDS epidemic in India.

In the United States and other economically advanced, democratic societies confronting the AIDS epidemic, a broad voluntarist consensus emerged from the early debates about HIV testing. This perspective drew upon American constitutional traditions and the liberal values that have informed those concerned with medical ethics, as well as upon pragmatic considerations. In part because of the United States' unique and early involvement with the AIDS epidemic and the compatibility of American constitutional values and international human rights standards, this perspective was to have a significant impact on the posture of the World Health Organization's Global Programme on AIDS.

Except for clearly circumscribed circumstances, testing was to be done under conditions of voluntary, informed consent, and the results were to be protected by stringent confidentiality safeguards. In the United States, to underscore the importance of protecting the privacy of tested individuals, the option of anonymity in testing centres was made broadly available.

Screening and behavioural change

Even before the screening of the blood supply had commenced in mid-1985, it was clear that the new antibody test would be enlisted in the effort to achieve the overriding public health goal of mass behavioural change. What was a matter of deep dispute and what has remained a matter of controversy was how much of a role testing would play. For public health officials who faced the daunting challenge of affecting the most intimate behaviours and for whom the legacy of other health promotion campaigns gave little reason for optimism, the test seemed to be a technology that could motivate the desired changes. Those who tested positive could be counselled about the urgent need for behavioural changes

to prevent the spread of HIV. Those who tested negative could be counselled about the importance of self-protection. Opponents of too great a reliance on testing – often representatives of the gay community and of civil liberties groups – were fearful of the social consequences of being identified as infected with HIV. Mass education and individual counselling, they argued, were the crucial options. From this perspective, testing, with all of its technical uncertainties, was too costly and too dangerous.

Over the years the bitter edge has all but vanished from the dispute. Public health officials have recognized that extraordinary changes have taken place in the behaviour of gay men, changes that could not be directly linked to the availability of testing. Such changes have, however, not been uniform. Mass education campaigns have not affected the sexual behaviours of some gay men and many intravenous drug users. What remains a matter of controversy is how aggressively to press for testing among those groups where the degree of behavioural change thus far attained gives cause for concern, or even alarm. Most recently, officials at the Centres for Disease Control have begun to acknowledge that the substantial resources committed to testing and counselling might not be the most effective approach for those at greatest risk.

Contributing to the always fragile consensus on testing for behavioural change has been a recognition on the part of public health officials of the need to win the confidence of those most at risk for HIV infection. To the fears that the confidentiality of test results would be breached, thus exposing those infected to stigma and discrimination, public health officials have responded by embracing the cause of confidentiality and legislation to protect the social and economic rights of individuals with HIV infection[6–9]. Without careful assurances of confidentiality and protection against irrational discrimination, encouraging those at risk to come forward for testing would be impossible.

Screening for clinical purposes

Early in the history of the epidemic, when medicine was all but impotent against the opportunistic diseases that afflicted those with HIV, the question with which some patients and many of their advocates confronted those who proposed antibody testing was: 'Of what benefit will a positive finding be for *me*?' Further, this was not a test like others, that clinicians generally used and that were typically covered by the broad, general consent implied in an agreement to be treated. Rather, critics argued, the antibody test was for social and psychological reasons, more like those invasive procedures for which special consent was required. Faced with such challenges and the undeniable reality of social stigma associated with HIV infection, physicians, their professional associations, and public health officials agreed that an exacting standard of consent to HIV testing was appropriate: specific informed consent was to be sought from patients or their surrogates. To many clinicians, however, such requirements represented an unacceptable intrusion into the therapeutic relationship, a hurdle designed to impede sound diagnostic work.

Now that medicine has made some remarkable advances in its capacity to control the life-threatening illnesses HIV-infected individuals face, given the increasing frequency with which physicians and researchers have begun to promote the prophylactic use of therapeutic agents and given the range of clinical trials for which infected people may be eligible, the picture for patients is very different from what it had been early in the epidemic. Under such circumstances one can answer the question, 'Of what benefit will a positive finding be for *me*?' quite differently than in 1985.

Under such circumstances, clinicians have begun to exert pressure to loosen the requirements for specific, informed consent before testing. In short, they have begun to maintain that the time has come to return AIDS to the medical mainstream[3].

Although one can appreciate clinicians' impatience as expressed by this perspective, resisting the pressure for routine testing without consent is important. In the first place, as much as the clinical picture has begun to change, no definitive therapeutic course is available for people infected with HIV but who are asymptomatic. At the same time, the possibility of stigma and discrimination remains an ever-present threat to the social well-being of the infected. Given these circumstances, arguments for specific, informed consent remain as pertinent as ever[10]. But even if the clinical picture improved dramatically, the moral grounds for insisting on informed consent before HIV testing would not change. It is an established principle of medical ethics that competent adults have the right to decide whether or not to undergo treatment and whether or not to terminate treatments already begun.

Certainly, the principle that limits the physician's paternalistic authority to order therapies in the interest of the patient extends to the authority to order tests that would serve as the basis for commencing treatment. But if respect for the autonomy of the patient requires that physicians exercise restraint, then their ethical responsibility to provide appropriate care now requires that they routinely offer HIV antibody testing to those patients whose social histories suggest an increased possibility of infection and for whom therapy is available and accessible.

Screening for safety

Because AIDS represented the first major infectious disease with which advanced industrial societies had had

to contend in almost a generation, and because the causative agent was not identified until three years after the first cases had been reported, it is not surprising that it provoked considerable social anxiety. Employers, landlords, schools and even some health care institutions evidenced a willingness to exclude those with the new disease. With the discovery of HIV and the development of a test that could detect the antibody to the virus, the potential for discriminatory activity increased despite the epidemiological evidence about how infection is spread.

In the United States, the Centers for Disease Control moved swiftly to contain the irrational impulse toward exclusion[11–13]. The message was clear: HIV could not be transmitted casually and thus no grounds existed for excluding infected individuals who were otherwise capable of performing their expected functions, from their schools or workplaces. In the context of the health care, setting universal blood and body fluid precautions would protect workers not only from HIV but also from the far more infectious hepatitis B. Screening could only provide illusory protection.

Although shameful efforts to exclude or isolate school children with AIDS or HIV infection and cases of discrimination by employers still occur, they are almost universally deplored as irrational, unscientific and retrograde.

The situation in health care has not been so clear-cut. The relatively few cases of transmission as a result of needle sticks[14], and the even smaller number of cases of transmission linked to blood splashes have provoked distress among health care workers, especially among surgeons, obstetricians, nurses and emergency room personnel. Those whose work regularly brings them into contact with their patients' blood feel vulnerable. On rare occasions they have asserted the right to refuse to care for those infected with HIV, thereby rejecting the fundamental principles of medical ethics[15, 16]. Far more frequently they have publicly challenged the adequacy of the recommendations for universal blood and body fluid precautions. They have demanded the right to know whether or not their patients are infected and the right to screen on a routine or mandatory basis for HIV infection. The level of vigilance demanded by the threat of a lethal infection cannot, they have argued, be maintained at all times. Protection requires knowledge of infection. The conflict between public health officials who have declared that universal precautions are adequate and clinicians who have asserted that they are not has not abated.

Haunting the entire debate about screening patients for safety reasons is the spectre of the refusal to treat. Those identified as infected will not simply be treated with greater vigilance. They will be given different care, lesser care, perhaps no care. In the absence of ironclad assurances to the contrary, those committed to protecting the infected will view all calls for mandatory screening of patients as a dangerous first step.

In the coming years, new controversies about screening will surely arise. In each conflict those who confront each other will predictably seek to appropriate the mantle of value-free decision making and will charge that those with whom they disagree have deserted the standards of science. On some occasions, the risk posed by the infected will be understood to be so small and the implications of screening and exclusion so burdensome that even the most cautious will find justifying compulsory testing and imposing restrictions hard. On other occasions the choices will not be clear-cut. Nevertheless, in each case more than 'science' will be involved. Decisions about screening policy will reflect the balance of moral commitments to privacy, reason and communal well-being.

Seroprevalence studies

With the identification of HIV as the aetiological agent of AIDS in 1984, investigators realized that an understanding of the dimensions of the epidemic would require more than a careful detailing of the incidence and prevalence of overt symptoms of the disease. To enable public health authorities to target and evaluate preventive interventions and plan for the health care services that would be required in the future as those infected progressed to symptomatic conditions, knowledge of the incidence and prevalence of HIV infection was critically needed.

Soon after antibody testing became possible, epidemiological specialists and public health officials concerned with the surveillance of the HIV epidemic realized that data based on volunteer studies involving only consenting individuals drawn from high-risk groups such as homosexual men, intravenous drug users, visitors to clinics for sexually transmitted diseases or from low-risk groups were insufficient for monitoring the incidence and prevalence of HIV infection because of selection and participation biases.

Anonymous, blinded or unlinked screening emerged as the method of choice for epidemiologists who sought to overcome the inadequacies of volunteer studies. Such screening would involve blood specimens collected for purposes other than HIV testing under conditions that permanently stripped such samples of personal identifiers. As an epidemiological tool, blinded studies had been used on other occasions to establish vaccination levels in populations and for reference purposes in hospital laboratories.

Because individuals participating in blinded screening could not be identified, researchers in the United States generally believed that the principles of privacy and informed consent were not being violated[17]. However, the very element of blinded studies that appeared to permit testing without negating these principles raised

problems for public health, and seemed to represent a breach of the ethical duty to inform individuals about clinical findings germane to their well-being. In the United Kingdom a leading medical ethicist based his profound objection to blinded studies on these grounds[18].

> We trade on deceit – a minor deceit but undoubtedly a deceit – if without either explicit or implicit permission we start using our patients for the benefit of others. The deceit is compounded if in so using our patients we discover important information that they may wish to know and we have deliberately both failed to find out whether or not they would wish to know it and so organized matters that we cannot pass it on even if they did wish to know.

Such objections were persuasive for a time in Great Britain, and in the Netherlands[19] and Denmark[20], where people have remained deeply suspicious of blinded studies. In the United States, however, such studies proceeded with virtually no dispute.

Those who propose to undertake such studies must now pay attention to the prospects of effective early intervention for those with HIV infection. The issue is no longer one of balancing the public health benefits of blinded studies against the potential benefits of studies that would not preclude the notification of infected people. Those with HIV infection have a clear and immediate interest in knowing that they are infected. The ethical standards that impose upon clinicians a duty to inform their patients may increasingly seem to be at odds with the ethical and professional duty of public health officials to develop the most accurate epidemiological foundations possible for guiding preventive interventions and organizing health care services. Under these circumstances, ensuring that the subjects of blinded seroprevalence studies have access to voluntary, confidential HIV antibody testing with appropriate clinical follow-up will be critically important.

Conclusions

India is beginning to experience a striking and disastrous increase in HIV infection. This epidemiological threat is taking place in a vast and populous nation with limited resources compared to those available to the industrialized democracies. The fundamental ethical challenges presented by the AIDS epidemic in India will basically be no different from those in the United States: to inhibit the spread of HIV infection, to provide medical care to the level permitted by the resources available, and to treat those individuals infected with HIV or diagnosed with AIDS with dignity and respect.

It is within the context of these ethical challenges that India must confront the question of how best to employ the HIV test. Like any other technology it can be used to the advantage of those who are vulnerable or it can be abused, thereby further burdening those most in need of protection, care and respect. How India chooses to employ the HIV test will represent an important measure of its commitment to both the public health needs imposed by the AIDS epidemic and to the human rights that may be compromised in the face of the challenge of a growing epidemic.

1. Bayer, R., *Private Acts, Social Consequences: AIDS and the Politics of Public Health*, Rutgers University Press, New Brunswick, NJ, 1991.
2. Mathan, V. I., The management of HIV infection in India, *Indian J. Med. Res.*, 1993, **97**, 225–226.
3. Bayer, R., Public Health Policy and the AIDS Epidemic: An End to HIV Exceptionalism? *N. Engl. J. Med.*, 1991, **324**, 1500–1504.
4. Narain, J. P., Pattanayak, S. and Shah. N. K., HIV testing policies: An overview, *Indian J. Med. Res.*, 1993, **97**, 219–222.
5. Lal, Shiv, and Thakur, B. B., Proposed national HIV testing policy, *Indian J. Med. Res.*, 1993, **97**, 223–224.
6. Centres for Disease Control, Recommended additional guidelines for HIV antibody counselling and testing in the prevention of HIV infection and AIDS, Atlanta, Georgia, 1987.
7. Surgeon General, Report on Acquired Immune Deficiency Syndrome, Washington DC, 1986.
8. Institute of Medicine and the National Academy of Sciences, *Confronting AIDS*, National Academy Press, Washington DC, 1986.
9. Presidential Commission, Report of the Presidential Commission on the Human Immunodeficiency Virus Epidemic, Washington DC, 1988.
10. Levine, C., and Bayer, R., The ethics of screening for early intervention in HIV disease, *Am. J. Publ. Health*, 1989, **79**, 1661–1667.
11. Centres for Disease Control, Education and foster care for children infected with human T-lymphotropic virus Type III/lymphadenopathy-associated virus, *Morb. Mortal. Weekly Rep.*, 1985, **34**, 517–521.
12. Centres for Disease Control, Recommendations for preventing transmission of infection with human T-lymphotropic virus type III/lymphadenopathy-associated virus, *Morb. Mortal. Weekly Rep.*, 1985, November 15, 681–686, 691–695.
13. Centres for Disease Control, Recommendations for preventing transmission of infection with human T-lymphotropic virus type III/lymphadenopathy-associated virus during invasive procedures, *Morb. Mortal. Weekly Rep.*, 1986, April 11, 221–223.
14. Centres for Disease Control, Update: Acquired immunodeficiency syndrome and human immunodeficiency virus infection among health-care workers, *Morb. Mortal. Weekly Rep.*, 1988, **37**, 229–239.
15. *New York Times*, 9 November 1987.
16. Zuger, A. and Miles, S. H., Physicians, AIDS and the occupational risk: Historic traditions and ethical obligations, *J. Am. Med. Assoc.*, 1987, **258**, 1924–1928.
17. Bayer, R., Lumey, L. H. and Wan, L., The American, British and Dutch responses to unlinked anonymous HIV seroprevalence studies: An international comparison, *AIDS (UK)*, reprinted in *Law, Medicine & Health Care*, 1991, **19**, 3–4.
18. Gillon, R., Testing for HIV without permission, *Br. Med. J.*, 1987, **294**, 281–283.
19. van Wijngaarden, J. L., The Netherlands: AIDS in a consensual society, in *AIDS in the Industrialized Democracies: Passions, Politics, and Policies*, (eds Kirp, D. L. and Bayer, R.), Rutgers University Press, New Brunswick, NJ, 1992.
20. Albaek, E., Denmark: AIDS and the Political Pink Triangle, in *AIDS in the Industrialized Democracies: Passions, Politics and Policies* (eds Kirp, D. L. and Bayer, R.), Rutgers University Press, New Brunswick, NJ, 1992.

The AIDS epidemic and economic policy analysis

David E. Bloom* and Ajay Mahal**

*Department of Economics, Columbia University, New York, NY 10027, USA, and
*Harvard Institute for International Development, 1 Eliot Street, Cambridge, MA 01238, USA
**Department of Economics, Barnard College, Columbia University, 3009 Broadway, New York, NY 10027, USA

ECONOMISTS have a vital role to play in helping public health officials and policy-makers understand the AIDS epidemic and design efficient policies to limit its impact. AIDS is first and foremost a public health problem, but it is a problem with deep economic roots and potentially devastating economic consequences. The main purpose of this article is to document this assertion.

The costs of AIDS

To understand the economic consequences of AIDS, a useful starting point is to examine the costs associated with this disease. These costs fall into four main categories. First are the personal medical care costs associated with AIDS, which capture the costs of detecting, treating and caring for people with AIDS. Second are the nonpersonal costs, such as the costs of blood screening, the costs of information, education and communication, and the costs of basic research on AIDS. Third is the cost of lost output and lost income because of AIDS morbidity and mortality. Fourth are the psychological costs associated with the epidemic, such as the pain and suffering caused by AIDS and the cost imposed upon people who must behave differently to avoid contracting or transmitting HIV.

The first two categories of AIDS costs jointly make up the direct costs of the epidemic, so named because they involve a direct diversion of scarce resources from other uses. The third category is referred to as the indirect cost of the epidemic. It is not an out-of-pocket cost like medical expenditures on AIDS patients. Rather, it is what economists term an opportunity cost, because it reflects the value of lost opportunities to work and produce the goods and services whose consumption yields satisfaction. Indirect costs are just as important a component of the total cost of the epidemic as direct medical costs. The epidemic's psychological costs are rarely discussed or quantified.

The first generation of economic research on the AIDS epidemic has involved measuring the direct costs of HIV and AIDS and the income foregone because of AIDS morbidity and mortality in many countries throughout the world. Every study of this issue of which we are aware concludes that AIDS is an unusually costly illness on a per case basis, for two main reasons. First, many of the opportunistic infections with which AIDS is definitionally associated, such as tuberculosis, pneumonia and cryptococcal meningitis, are costly to treat relative to the cost of treating other common illnesses; and second, AIDS disproportionately afflicts individuals in their prime productive years, thereby causing a considerable loss of income to them and their families.

Economists have estimated that the lifetime personal medical care costs associated with AIDS are more than double India's per capita income, assuming that people with AIDS receive medical care that conforms fully to local protocols for the treatment of AIDS and the opportunistic infections with which AIDS is associated. The income lost because of AIDS morbidity and mortality is estimated to be far greater: more than ten times India's income per capita, in discounted terms[1]. This estimate of the loss of private income should not be confused with social output losses, especially in a labour surplus setting, although the two concepts can be linked. Overall, these estimates of medical care costs and lost income indicate that each case of AIDS is quite costly in India, just as it is elsewhere in the world.

Economic determinants of HIV transmission

The contribution of economics to understanding the AIDS epidemic is not limited to quantifying its impact in monetary terms. Economics also provides a powerful way of examining the pattern of the epidemic's spread. The central idea is that HIV is not spread randomly, as tends to be the case with the bacteria that cause tuberculosis or the virus that causes the common cold. Rather, HIV is most often transmitted as a consequence of purposeful behaviour that often has a strong economic foundation. For example, some of the sex that contributes to the spread of HIV is sold on the street and through brothels[2,3]. As another example, a close connection seems to exist between labour migration and HIV transmission[4–7]. This connection seems to arise because the behaviour of migrants is not easily subject to monitoring by their families and communities, because migrants often have considerable amounts of cash to spend, and because they may be lonely because of being separated from their families. In addition, because migrants are often poorly educated and not closely con-

nected to stable communities, making them aware of the behaviours that put them at high risk for contracting or transmitting the virus that causes AIDS can be challenging. All these factors increase the likelihood of migrants engaging in commercial sex, an activity that puts them at high risk for contracting HIV, and ultimately places their spouses and unborn children at risk.

India has a huge migrant population, both internal and international. For example, rural to urban migration is significant, especially in Maharashtra, Tamil Nadu and West Bengal. So too is the floating population of migrant workers[8]. India is also experiencing substantial cross-border population movements with Bangladesh, Nepal, and to some extent Myanmar, as well as strong migration flows to and from the Middle East and a growing influx of foreign tourists, business people, and others. Add to this a continuing process of economic reform, which is likely to promote the movement of people into, out of, and within India, and the situation is one of continuing and increasing vulnerability to the spread of HIV, especially among the general population.

We conclude from this reasoning and these examples that economists have a key role to play in understanding the nature of the HIV and AIDS epidemic and in assessing its impact. Economics is a crucial ingredient in any explanation of why HIV spreads so rapidly among lorry drivers, fishermen, sailors, construction workers, overseas contract workers, the military, commercial sex workers, injecting drug users, and commercial blood donors. That so many countries throughout the world share this pattern of spread is no coincidence.

Does government have a role to play?

Government intervention in the area of HIV/AIDS encompasses a myriad of activities aimed both at controlling the spread of HIV and coping with the immediate shock of the epidemic and its wide-ranging echo effects. Interventions may be justified on the grounds that they promote efficiency (that is, increase the size of the economic pie) or promote equity (achieve a more desirable distribution of the pie).

Intervention can promote efficiency if one individual's behaviour imposes costs upon others for which no compensation is paid, a market failure that economists refer to as a negative externality. As an example, an externality may arise if a person with AIDS and tuberculosis coughs or sneezes in public, and thereby subjects other nearby individuals to the risk of tuberculosis infection. Costs that are not easily deterred are being imposed without offsetting compensation. As another example, consider women who are powerless to negotiate the use of condoms with their sex partners, a situation that arises in many commercial sex establishments as well as in many marital relationships. To the extent these women are exposed, along with their unborn children, to the risk of HIV infection, costs that are not easily deterred are again imposed without compensation, a classic market failure for which government intervention is a natural remedy.

Intervention may also promote efficiency if private individuals are poorly informed about the risks and costs of HIV transmission. Indeed, under certain conditions the government can improve well-being either by providing information about HIV and AIDS or by adopting tax, subsidy, or regulatory policies that induce people to act as if they had this information.

An equity justification for intervention could arise if the costs of the epidemic fall disproportionately upon the poor[1,9]. If the government deems the ensuing distribution of economic well-being, under which the poor are further immiserized, to be socially undesirable, government intervention aimed either at preventing HIV transmission or redistributing resources is well justified.

Contribution of economic reasoning to policy-making

Even though making a case for some government intervention in the area of HIV and AIDS in Asia's developing countries is not difficult, the complexity of the mechanisms that lead to HIV infection and that determine their social and economic impacts make effective policy-making an especially challenging task. We offer three examples to illustrate this point.

The first example relates to HIV transmission via the sharing of needles among injecting drug users. To reduce the rate of HIV transmission, some policy-makers have suggested reducing the number of drug users by cracking down on the supply of opium. From an economist's perspective, however, destroying opium fields might actually promote HIV transmission, a result exactly opposite to the one desired. This could happen if a reduced supply of opium leads to an increase in its price, which in turn leads habitual opium smokers to cut back on their opium consumption by switching to injection, a far more efficacious mode of opium use[10].

Throughout North America and Western Europe, policy-makers are increasingly responding to the problem of HIV transmission through the sharing of nonsterile injection equipment by establishing needle exchange programmes[11]. Under these programmes, injecting drug users can receive new injection equipment at no charge in exchange for used injection equipment. Needle exchange programmes are controversial because they try to improve the conditions under which illegal drug use takes place, which critics claim promotes greater drug use. Nonetheless, preliminary evidence for the United States suggests that the rate of new HIV infection among injecting drug users who participate in such programmes is lower than the corresponding rate among nonparticipants[12].

As a second example, consider a policy of isolating or imprisoning seropositive commercial sex workers, as India did for a brief period several years ago in a highly publicized case in which 854 commercial sex workers, 457 of whom were HIV-positive, were moved from Bombay to Madras and confined[3]. On the surface, this policy might seem to be a sensible way to curb HIV transmission, because it prevents seropositive sex workers from selling sex. However, it fails to address the demand side of the market for sex, that is, the fact that certain individuals, be they foreign businessmen or native sailors, fishermen or migrant workers, will continue to demand and be willing to pay for sex. Under the economic conditions that currently prevail throughout most of the world, that demand will be satisfied by a new group of women who will supply sex commercially, not because they are social deviants, but because they lack better economic opportunities. Although these women may be seronegative when they enter the sex industry, without information about HIV and without the power to negotiate the use of condoms, they too will become infected and then infect others.

A third example concerns HIV testing. Some policymakers believe that mandatory testing is a useful approach to HIV prevention[13]. Their rationale is that people will respond altruistically to the knowledge that they are HIV-infected by curbing behaviours that put others at high risk of infection, especially if they receive appropriate counselling. Altruism is, however, not the rule in economic behaviour. Selfishness is much more common, and in this context might actually cause HIV testing to promote further transmission as people who learn that they are seropositive recognize that they no longer face a key disincentive to high-risk behaviour: the risk of HIV infection. Although research findings vary, a number of reputable studies have found evidence that positive HIV test results trigger increases in high-risk behaviour, thereby supporting the view that testing may promote HIV transmission[14,15].

These three examples are not meant to suggest that economic reasoning provides a complete explanation of individual decisions about drug consumption, the supply of and the demand for sex, or other high-risk behaviours. They do, however, illustrate one of the hallmarks of economists' approach to policy-making: the need to understand the mechanism that is generating undesirable outcomes as a prelude to effective policy design and implementation.

In the first example, the economics of drug injection suggests that a policy of destroying opium fields might increase drug injection more than it decreases drug use, the net effect being a higher rate of HIV transmission. This reasoning does not, of course, imply that the government should promote opium production as a way to drive prices down and encourage smoking in place of injection. While such a policy might, in principle, curtail the spread of HIV, it would probably do so quite slowly, because heavy drug injectors rarely revert to smoking. More important, it might simultaneously increase the number of drug users, which is clearly not a desirable outcome. However, this bit of economic reasoning does help explain the calls for decriminalizing the sale, possession and use of drugs that one increasingly hears in North America and Western Europe. It also highlights the fact that many practical policies for HIV prevention may conflict with other social goals and policies. For example, needle exchange programmes may increase the number of drug users, early sex education and free contraceptives may encourage promiscuity, condom distribution in the prison system may condone illicit homosexual activity, and policies mandating the use of condoms in brothels may undermine the illegality of prostitution.

In the second example, a policy of imprisoning seropositive commercial sex workers is little more than a temporary solution, what might best be termed a band-aid solution, to the problem of HIV transmission. It ignores both the demand side of the market for sex and the readily available supply of additional sex workers.

In the third example, as already noted, mandatory testing may actually promote HIV transmission. This observation highlights the fact that the HIV epidemic can, in principle, be controlled without identifying a single HIV-infected person. As long as the population behaves carefully, a goal that is presumably promoted by the provision of information, those who are infected will not transmit the HIV and those who are uninfected will not contract it. Testing the blood supply and doing some testing for purposes of epidemiological surveillance may be desirable, but this testing can be done anonymously. Limiting the amount of testing that is done also affords the greatest protection to individual privacy, thereby limiting the ostracism and stigmatization that so often, so needlessly, and so counterproductively accompany the spread of HIV and AIDS.

Contribution of economic tools to policy-making

Good economic reasoning encourages policy-making directed at underlying problems, not at superficial symptoms. Economics also offers a set of tools that are often extremely useful for rational policy-making. These tools, known as cost–benefit analysis and cost-effectiveness analysis, are methods for evaluating the desirability of alternative policy options[16]. Both methods require us to estimate the costs of implementing alternative policies. Cost–benefit analysis also estimates the expected benefits corresponding to each alternative, and the optimal policy is the one that has the highest net benefit, that is, benefits less costs.

Cost-effectiveness analysis is closely related to cost–benefit analysis, except that it stops short of estimating

the monetary benefits of particular policies. Instead, it calculates and compares the cost of alternative methods of achieving a particular qualitative outcome.

As a brief illustration, suppose we wish to compare the following eight programmes for HIV prevention and control: (1) mass education through national television and radio; (2) mass education through the distribution of a one-page AIDS information sheet to every household in India; (3) a special HIV and AIDS information programme for Indian workers and tourists going abroad; (4) a special HIV and AIDS information programme for lorry drivers; (5) a programme to expand the provision of family planning services to unmarried individuals; (6) a programme to treat sexually transmitted diseases (STDs), which is important because STDs greatly increase the risk of HIV transmission; (7) a programme for screening India's blood supply for HIV; and (8) a community-based programme aimed at reducing the demand for drugs and promoting harm reduction among injecting drug users in Manipur and other areas with relatively heavy drug use. Using cost–benefit analysis we would estimate the number of HIV infections that each programme would avert and place a monetary value on those averted infections using estimates of the direct and indirect costs of the epidemic. We would then compare the estimated monetary value of averted infections and of any other expected consequences of each programme, such as control of STDs, fertility reduction, reduced drug use, lower crime rates, and so on, with the cost of each programme. Good programmes are those whose benefits exceed their costs, much like good investments are those that yield positive net returns.

Using cost-effectiveness analysis, we would not place a monetary value on the estimated number of averted HIV infections. We would simply calculate the cost per averted infection under each programme as a way to identify which of them offers the least expensive way to control the spread of HIV. Unlike cost–benefit analysis, which takes multiple programme impacts into account and indicates whether each programme is a good investment or not, cost-effectiveness analysis simply ranks alternative programmes in terms of their economic efficiency in achieving a particular outcome. As cost-effectiveness analysis does not require the monetization of programme benefits, it also requires fewer assumptions and calculations than cost–benefit analysis.

Cost–benefit and cost-effectiveness analyses provide a systematic, sensible and practical way of selecting among alternative policy options. Combinations of options may also be analysed, which is important if programmes complement each other. The analyses may be performed from the point of view of an individual, an organization, or society as a whole, and often yield different results depending upon whose point of view is adopted. Although these analyses are sometimes difficult to carry out because of data limitations and a variety of difficulties inherent in calculating and monetizing programme impacts, this is not always the case, as some recent cost–benefit analyses for China and Sri Lanka on the costs and benefits of screening the blood supply for HIV demonstrate[7,17].

Considerable work remains to be done in applying the tools of cost–benefit and cost-effectiveness analyses to the evaluation of policies designed to promote HIV prevention and control. Especially interesting would be analyses of HIV and STD prevention programmes that involve expanding the distribution of condoms in India and of programmes for both mass and targeted education about HIV and AIDS, especially programmes designed to reach migrants. In the absence of a cure for HIV and AIDS, information about safe behaviour and the necessary motivation to practise it are the only options even remotely resembling a vaccine against HIV transmission. HIV can spread fast, but so too can information. Indeed, information delivered through community-based programmes and channels of influence appears to be the most potent ally available in the effort to control further spread of HIV[18].

Impact of AIDS on economic growth and development

Leaders in the global campaign against AIDS have repeatedly expressed the opinion that the AIDS epidemic is likely to have a substantial – perhaps even a catastrophic – impact on the macroeconomies of developing countries[15, 19–22]. A variety of economic studies support these assertions. These studies indicate that in the hardest hit countries of Africa, the HIV and AIDS epidemic may diminish the annual growth rates of national income by as much as two-thirds of a percentage point, and diminish the annual growth rate of national income per person by as much as one-third of a percentage point through the year 2010 (refs 23–26). In relation to the growth rates of many of the world's economies, two-thirds of a percentage point is not as little as it may seem initially. Also, it adds up over time, which is important given that a biochemical vaccine against HIV infection and a cure for AIDS are nowhere in sight.

The economic simulation models and cost of illness calculations that support the prediction that AIDS will have dire macroeconomic consequences require many assumptions that are difficult to verify, but adopting a more direct approach to examining the impact of the epidemic on the growth rates of income per person is now feasible. We are in the second decade of the world AIDS epidemic, which has progressed relatively far in several countries, especially in Africa. For example, nearly 20% of the adult populations of Uganda and Zambia are infected with HIV. In Zambia, the cumulative number of cases of full-blown AIDS has already

reached 4% of the adult population. With this range of HIV experience, from low prevalence countries like China and Sri Lanka, to medium prevalence countries like India, Myanmar and Thailand, to high prevalence countries like Uganda and Zambia, we may examine whether countries that have experienced more severe AIDS epidemics have also experienced lower rates of economic growth, while controlling for other influences on economic growth and addressing the possible reciprocal influence of economic growth on the spread of the HIV.

Technical issues and other details aside, preliminary calculations based on data for 51 countries reveal that the AIDS epidemic has had only a small and statistically insignificant negative effect on income per person. These results resemble ones derived through similar analyses of the economic impacts of the world influenza epidemic of 1918–19 and the Black Death of the mid-1300s, both of which killed massive numbers of people in extremely short periods of time[28].

With these findings in hand, and given that India already has first-generation estimates of the economy-wide costs of the epidemic, the amount of attention devoted to the macroeconomic impacts of the AIDS epidemic can be reduced. This will give policy-makers and researchers an opportunity to focus greater attention on the impact of the epidemic on particular industrial and occupational sectors, on particular geographic regions, and on particular demographic groups, especially people living with AIDS and their families, women, the poor and other marginalized groups, such as homosexuals and drug users. These are the areas that health policymakers, economists, international organizations, nongovernmental organizations, and others concerned with the social and economic implications of the epidemic need to address most urgently. They involve complex issues that include matters of health, sociology, psychology, law, politics and economics. As research problems, they call for multidisciplinary approaches, mirroring their nature as multisectoral problems from the standpoint of policy design and implementation.

1. Bloom, D. and Glied S., Who is bearing the cost of the AIDS epidemic in Asia? in *Economic Implications of AIDS in Asia* (eds Bloom D. and Lyons, J.), United Nations Development Programme, 1993.
2. Centre for International Research, *HIV/AIDS Surveillance Data Base*. United States Bureau of the Census, Washington DC, 1993.
3. Nataraj, S., Madras, India: Locking up prostitutes, in *The Third Epidemic: Repercussions of the Fear of AIDS*, The Panos Institute, London, 1991.
4. Bloom, D. and Mahal, A., Economic Implications of AIDS in Asia, Draft, 1995.
5. Gharpure, H., Chanderkar, N. and Sengupta, S., Serosurveillance of HIV Infection in Migrant Workers: Socioeconomic Relationship, Abstract No. B314 at the Second International Conference on AIDS in Asia and the Pacific, New Delhi, 1992, November 8–12.
6. Solon, O. and Barrozo, A., Overseas contract workers and the economic consequences of the HIV/AIDS in the Philippines, in *Economic Implications of AIDS in Asia* (eds Bloom, D. and Lyons, J.), United Nations Development Programme, New York, 1993.
7. United Nations Development Programme, The Socioeconomic Dimensions of the HIV/AIDS Epidemic in Sri Lanka: A Report, Colombo, Sri Lanka, 1995.
8. United Nations Development Programme, HIV/AIDS Situation in India, Country report, presented at the UNDP/Asian Development Bank Workshop on the Economic Implications of the HIV/AIDS Epidemic in Selected DMCs, Manila, Philippines, 1993, September 5-8.
9. Stiglitz, J., *Economics of the Public Sector*, W. W. Norton, New York, 1988, 2nd Edition.
10. Strang, J., Des Jarlais, D., Griffiths, P. and Gossop, M., The study of transitions in the route of drug use, the Route from one route to another, *Br. J. Addiction*, 1992, **87**, 473–483.
11. Lurie, P. and Reingold, A., *The Public Health Impact of Needle Exchange Programmes in the United States and Abroad*, University of California, Berkeley, 1993.
12. Des Jarlais, D., Friedman, S., Sotheran, J., Wenston, J., Marmor, M., Yancovitz, S., Frank, B., Beatrice, S. and Mildvan, D., Continuity and change within an HIV epidemic: Injecting drug Users in New York City, 1984 through 1992, *J. Am. Med. Assoc.*, 1994, **271**, 121–127.
13. Presidential Commission on the Human Immunodeficiency Virus Epidemic, Report, Washington DC, 1988.
14. Bloom, D. and Glied, S., Projecting the number of new AIDS cases in the United States, *Int. J. Forecasting*, 1992, **8**, 339–365.
15. Philipson, T. and Posner, R., *Private Choices and Public Health: The AIDS Epidemic in an Economic Perspective*, Harvard University Press, Cambridge, 1993.
16. Drummond, M., Stoddart, G. and Torrance, G., *Methods for the Economic Evaluation of Health Care Programmes*, Oxford University Press, New York, 1987.
17. Jianhua, Y., Jingyuan, Yu, Tao, J. and Yi, Xu, Blood screening in China: An economic analysis, Draft, Beijing Institute of Information and Control, Beijing, China, 1995.
18. Coates, T., Communities and AIDS prevention: What works, AIDS captions 1994, May 2–4.
19. Bloom, D. and Lyons, J. (eds), *Economic Implications of AIDS in Asia*, United Nations Development Programme, New York, 1993
20. Merson, M., The AIDS Epidemic in Asia: The Reality, the Opportunity, and the Challenge, Keynote address, delivered at the Second International Congress on AIDS in Asia and the Pacific, New Delhi, 1992, November 8–12.
21. Panos Institute, *The Hidden Cost of AIDS: The Challenge of HIV to Development*, Panos Publications, London, 1992.
22. United Nations Development Programme, The HIV Epidemic as a Development Issue, United Nations, New York, 1992.
23. Cuddington, J., The Macroeconomic Impact of AIDS in a Dualistic, Labor-Surplus Economy with an Application to Tanzania. Working Paper #92-12, Georgetown University, Washington DC, 1992.
24. Cuddington, J., Modeling the macroeconomic effects of AIDS, with an application to Tanzania, *The World Bank Economic Review*, 1993, **7**, 173–189.
25. Cuddington, J. and Hancock, J., forthcoming, Assessing the impact of AIDS on the growth path of the Malawian economy, *J. Develop. Econ.*
26. Over, Mead, The Macroeconomic Impact of AIDS in subSaharan Africa, Technical Working Paper No.3, The World Bank, Technical Department, Population, Health, and Nutrition Division, Washington DC, 1992.
27. Bloom, D. and Mahal, A., Does the AIDS Epidemic Really Threaten Economic Growth? Working Paper No. 5148, National Bureau of Economic Research, Cambridge, MA, USA, 1995.

Contributors to the Special Section

Ronald Bayer is a Professor at the Columbia University School of Public Health. His interests centre on justice and health care, occupational health, and controversies in science and technology. He has written extensively on ethical and political controversies in medicine.

David E. Bloom is a Professor of Economics and former Chair of the Department of Economics at Columbia University. During 1995–96, he is serving as Executive Director of the Harvard Institute for International Development (HIID). Since 1991, he has been a member of the board of reviewing editors of the magazine *Science*.

Tim Brown is currently a research Fellow in the Program on Population of the East-West Center in Honolulu, Hawaii. He served as a Research Associate with the Program on AIDS of the Thai Red Cross Society. His primary interests are in the areas of epidemiology, AIDS/HIV modelling, and risk behaviour assessment.

James Chin was the Chief of Surveillance, Forecasting, and Impact Assessment unit of the Global Programme on AIDS, WHO, from 1987 to mid-1992. He is currently a Clinical Professor of Epidemiology at the School of Public Health, University of California at Berkeley and an international consultant on the impact of AIDS in developing countries.

John M. Dwyer is the Chairman of the Department of Medicine at the Prince Henry/Prince of Wales Hospitals of the University of New South Wales and the Director of the AIDS Treatment and Research Center at those hospitals. In 1982, he started an AIDS Treatment and Research Center at Yale University. He is the founding President of the AIDS Society of Asia and the Pacific.

Sherry Glied is an Assistant Professor of Public Health and Economics at Columbia University and a faculty research fellow of the National Bureau of Economic Research in Cambridge, Massachusetts. She has done extensive research on various aspects of the AIDS epidemic in both the US, and India.

Shiv Lal is the Additional Project Director and Technical Head at India's National AIDS Control Organization (NACO). He has served as the Director of the National Malaria Eradication Programme. He holds degrees in epidemiology and community health.

Ajay Mahal is a visiting professor of Economics at Barnard College. He holds a Ph D in Economics from Columbia University. His dissertation was on the economic impact of AIDS.

George Petersen is the World Health Organization's representative in Cambodia. He has also served as the Regional Adviser in Communicable Diseases at the WHO's Regional Office for the Western Pacific.

Rabin M. Sarda is the Acting Regional Adviser on AIDS and STD at the Western Pacific Office of the World Health Organization. He is responsible for WHO support and collaboration with national AIDS programmes in 35 countries in the Western Pacific.

Werasit Sittitrai is currently serving as the Deputy Director of the Program on AIDS of the Thai Red Cross Society. He is a member of the WHO Steering Committee on Social and Behavioural Research on AIDS and the Thailand National AIDS Prevention and Control Committee.

B. B. Thakur is serving as the Assistant Director in the National AIDS Control Organization. Prior to joining NACO, he was the Deputy Assistant Director General of Community Health. He holds a degree in community medicine.

Nandrani S. de Zoysa served as the Director of Sri Lanka's National Blood Transfusion Services for 14 years.

Standard potential of hydrogen electrode

S. Parthasarathy, K. Ramya and V. K. Venkatesan

SPIC Science Foundation, 'Mount View', II Floor, 110, Mount Road, Madras 600 032, India

The basis for determination of standard potential of hydrogen electrode is discussed. The EMF values for five cells, each one having a half cell consisting of a hydrogen electrode with different activities/concentrations of hydrogen ions, the other half cell being either Ag/AgCl or An/AnCl$_2$ of different molalities, are measured using solid ion-transmitting bridge without liquid junctions to calculate thermodynamic activities. It has been concluded that the standard potential of hydrogen electrode is actually zero, as the measured EMF value for the cell Zn/ZnCl$_2$ (0.6299 m)/HCl (1.18 M)/H$_2$ (1 atm), Pt viz. –0.762 V is in accordance with the standard potential of zinc obtained from precise calorimetric data. The various factors that support this conclusion are presented.

RECENTLY it has been demonstrated that it is possible to measure single ion activities using a solid chloride ion transmitter bridge–AgCl bridge. Single ion activities of zinc and chloride in $ZnCl_2$ solutions in molality range 0.1–0.001 were reported[1].

Measurement of single ion activities should lead to determination of absolute electrode potentials. The assumption made for arriving at the standard potentials of various electrode system is that the standard potential of hydrogen electrode is zero. Precise calorimetric data for standard potential of zinc, which is not dependent on any potentiometric value, are available. It should, therefore, be possible to determine the standard potential of the hydrogen electrode using half cell potentials of zinc electrode reported earlier and the standard hydrogen electrode.

The hydrogen electrode used for the potential measurement was equipped with a presaturator, a hydrogen bypass and was designed to be completely immersed in thermostatted bath. The cell set up is described elsewhere[2]. The two hydrogen electrodes were platinized from a solution of 1% chloroplatinic acid containing 10 mg lead acetate with a current density of 30 mA/cm^2 for 10 min. The electrodes were then rinsed with distilled water several times.

For the preparation of solid-ion transmitting bridge, pure AgCl was melted into the bridge in the form of a 'U' tube made from pyrex glass with a capillary of 2 mm bore diameter. The tubes are then heated to 380°C and held at that temperature for 10 h and then allowed to cool.

Table 1. Mean activity coefficient of HCl at 25°C

Molality	Activity coefficient	Molality	Activity coefficient
0.001	0.966	0.5	0.769
0.005	0.930	1.0	0.811
0.01	0.906	1.5	0.898
0.05	0.833	2.0	1.011
0.1	0.798	3.0	1.31

The hydrogen electrode was immersed in water in a thermostat set to 25°C. The thermostat is a Haake D8, Germany capable of controlling the temperature by ±0.1°C. Pure hydrogen is passed to the cell by Hydrogen Generator model No. 8200 of Packard Instrument Company, Canberra.

It is reported by Vogel[3] that a solution of 1.18 M corresponds to unit activity and is supported by literature values[4].

From Table 1, it can be seen that 1.0 m corresponds to an activity of 0.811 and 1.5 m corresponds to an activity of 1.347. So an activity of 1.00 corresponding to molarity of 1.18 m follows from interpolation. Molarity of 1.18 corresponds to an activity coefficient of 0.8475. The modified Debye–Hückel equation (Davies equation) which would give the above mentioned activity coefficient is:

$$-\log\gamma_{\pm} = 0.509 Z_+ Z_- \left\{\frac{\sqrt{I}}{1+\sqrt{I}} - 0.3221\right\}.$$

This equation is acceptable as the empirical constant 'C' for highly concentrated solutions can vary between 0.3 and 0.4 in the formula proposed by Davies[5] for activity coefficient as:

$$-\log\gamma_{\pm} = 0.509 Z_+ Z_- \left\{\frac{\sqrt{I}}{1+\sqrt{I}} - CI\right\}.$$

Hence we chose the value of 1.18 M for preparing a HCl solution of unit activity.

The exact molarity was determined by standard titration technique. The HCl solution of 0.01 m was prepared by diluting the standard volumetric solution supplied by E. Merck India Ltd. Zinc chloride solutions were prepared by dissolving the exact weights of pure zinc in 1:1 HCl, evaporating to dryness and adding the requisite amount of de-ionized water.

If the activity of a solution of zinc chloride is assumed to be 1,

$$a_{ZnCl_2} = 4m^3\gamma_{\pm}^3 = 1,$$

Robinson and Stokes[6] showed that if mean molality and the mean activity coefficient is 1, molality of zinc ion would be $m_{Zn^{2+}} = (4)^{-1/3} = 0.6299$ m. In this high molality range, neither the Debye–Hyckel equation nor the Guntelberg equation will be applicable. Only the em-

pirical equation proposed by Davies is applicable. The single ion activity of Zn^{2+} ion will be equal to 1 from the expression

$$a_{Zn}{}^{2+} = 4m_{Zn}{}^{2+}\gamma_{Zn}{}^{2+}.$$

Since $ZnCl_2$ solutions were prepared with extreme care, the molality variation is ruled out and we have prepared a solution 0.6299 m $ZnCl_2$ to provide for unit activity of zinc ion corresponding to the standard state. This corresponds to an empirical constant C well within the range of 0.2 to 0.3 in the Davies equation, as in the case of HCl solution. Zinc chloride solution of 0.1 m was also prepared by the above procedure.

The instrument used for potential measurement was ORION expandable ion analyser, model No. EA 940, from ORION Research Inc., USA which is capable of handling signals from high resistance sources with negligible IR drop.

The bridges are tested by interconnecting Tacussel Calomel electrode in saturated KCl and Tacussel AgAgCl electrode also in saturated KCl in two different beakers, i.e. the cell

Ag/AgCl, KCl(s)/bridge/KCl(s),Hg_2Cl_2,Hg

The potential was seen to be in keeping with the expected value of 45.6 mV[7].

The hydrogen electrodes were tested by observing the potential of the following cells:

1. Pt,H_2 (1 atm), HCl (0.01M)/AgCl, Ag
2. Pt,H_2 (1 atm), HCl (1.18M)/AgCl, Ag

The observed potentials were 464.8 mV and 222.4 mV (standard potential for Ag/AgCl electrode) respectively and were found to be in accordance with values reported elsewhere[8,9].

It was considered necessary to test the bridge when one of the solutions was acidic as it is well known that the mobility of protons is much higher compared to other ions. For this, the bridge was tested by verifying the potential of the cell

Pt,H_2 (1 atm),HCl(1.18M)/Bridge/HCl(1.18M),AgCl,Ag.

The observed potential agrees with the standard potential of 222.4 mV for the silver-silver chloride electrode.

Table 2. Measured potentials using hydrogen electrode at 25°C

Cell assembly	Potential mV
Pt,H_2(1 atm), HCl(0.01 m)/Bridge/KCl(s), AgCl, Ag	318.0
Pt,H_2(1 atm), HCl(0.01 m)/Bridge/$ZnCl_2$(0.1 m), AgCl, Ag	409.5
Zn, $ZnCl_2$(0.1 m)/Bridge/HCl (0.01 m), H_2 (1 atm), Pt	671.9
Pt,H_2 (1 atm), HCl (1.18 M)/Bridge/KCl(s), AgCl, Ag	199.0
Zn, $ZnCl_2$(0.6299 m)/Bridge/HCl (1.18 M), H_2 (1 atm), Pt	–762.0

In order to determine the absolute potential of the hydrogen electrode, the potential of the following cells were measured and the measured potentials are given in the Table 2.

From the potential measured for cell 1 of 318 mV, if we deduct the well-established reference-electrode potential of 199 mV, we arrive at Pt,H_2(1 atm), HCl(0.01 m) potential of the single electrode system as –119 mV. This half cell is coupled with Ag/AgCl system with 0.1 m $ZnCl_2$ using the chloride ion transmitter bridge in cell 2. The measured potential of cell 2 is 409.5 mV. Substracting the hydrogen electrode potential gives a value of 290.5 mV for Ag/AgCl, $ZnCl_2$(0.1 m) half cell, which is in accord with the earlier measurements reported[1].

The potential of cell 3 was measured and was found to be 671.9 mV., i.e. coupling of the same hydrogen electrode with zinc electrode in 0.1 m $ZnCl_2$ solution through the bridge. Addition of the hydrogen electrode potential gives the potential of the zinc half cell, Zn,$ZnCl_2$(0.1 m) as –791 mV which is in accordance with values reported earlier. These data confirm the consistency of the potential measurements made with solid-ion transmitting bridge in acidic systems.

The zinc system was chosen because it is perfectly reversible with respect to the cation, does not form unipositive ions, does not undergo self corrosion, is highly reproducible and precise calorimetric values are available for enthalpy changes[10–12] which when coupled with entropy data, can lead to the standard potential of zinc from methods other than potentiometric methods (See Figure 1). But the final confirmation can come only from direct measurement from a solution of H^+ ions of unit activity.

The potential of the hydrogen electrode system dipping into a solution of unit H^+ ion activity was connected through a solid-ion transmitting bridge to a saturated KCl solution in which Ag/AgCl electrode was dipped (see cell 4, Table 2). The measured potential of 199 mV is in keeping with the expected value.

When the electrode is again dipped into a solution of unit H^+ ion activity and is coupled to a $ZnCl_2$ solution in which the activity of zinc ion is kept unity (See cell 5 Table 2), the voltage observed is –762 mV which is in accordance with the value derived from calorimetric measurements.

The experimentally observed standard potential of zinc using a solution of unit activity with respect to a zinc ion and HCl solution of unit activity with respect to hydrogen ion is surprisingly in agreement with the calorimetrically determined standard potential of zinc (see Table 2). This shows that standard potential of hydrogen electrode is not an arbitrary zero but is actually zero.

The potential of standard hydrogen electrode is conventionally taken as zero at all temperatures, when the partial pressure of H_2 is one atmosphere and the activity

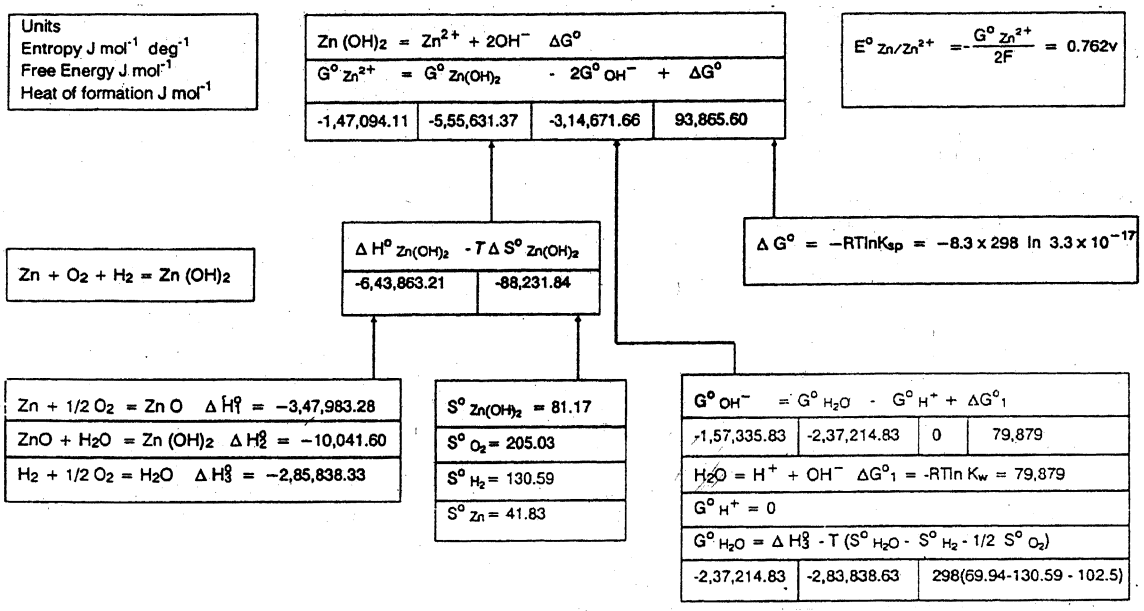


Figure 1. Standard potential of zinc.

of hydrogen is unity. The assumption that the $E^{o}_{H^+/H_2}$ is zero, leads to the consequent result viz. $\Delta G^\circ = 0$ which means that the free energy of reactant and product of the reaction

$$H^+ + e \rightleftharpoons 1/2\ H_2$$

is equal. The possibility of such a situation can be easily understood if one can take into consideration the following facts:

1. Proton, unlike other ions, is of Fermi dimensions (10^{-13} cm)–a Fermion. It is an elementary particle and is a nonconformist ion[13]. Classical theories are not applicable and quantum-mechanical considerations are possible.

2. The classical thermodynamic considerations are not applicable to fundamental particles like electron, proton etc. Therefore, the question of considering the entropy of the electron of the hydrogen atom which on de-electronation will get mingled with conductance or fermi lectrons of platinum, does not arise.

3. Proton's affinity for electron can be seen from the high ionization energy required i.e. 313 k cal.mole^{-1} for removal of an electron from hydrogen atom[14].

4. The heat of formation of hydronium ion H_3O^+ is 170 k cal.mole^{-1} which is considerably less than the ionization energy of hydrogen atom[15].

5. It is suggested that proton most probably exists in solution as H_3O^+ ions but the evidence is not all that certain[16].

6. Platinum serves not merely as an insert conductor but functions more as an efficient electrocatalyst[17] enabling the hydrogen molecule to dissociatively adsorb on the surface and help in the de-electronation of hydrogen atom and electronation of proton. The potential of platinum does not influence the hydrogen electrode potential. Large surface area can also be created[18].

7. The electrical double layer model of Bockris, Devanathan and Muller, which has been accepted as the one that explains many of the observed facts, envisages a layer of water molecules which are highly oriented and has probably a dielectric constant of 6–7 corresponding to a saturated dielectric[19]. The potential of zero charge for Pt in the presence of hydrogen in H_2SO_4 solution is reported[20] to be in the range + 0.11 to + 0.27 Vs standard hydrogen electrode. At the potential of zero charge, the water molecules are arranged with oxygen end and hydrogen end pointed towards the metal surface in an alternate manner to make the contributions of water molecule to the charge as zero. At the potential of standard hydrogen electrode, viz. zero which will be negative to the potential of zero charge, the charge on the Pt surface will be –ve and therefore the water molecules are adsorbed with their hydrogen ends toward the Pt surface. The presence of hydrogen ends adjacent to Pt surface will not enable hydrogen ion to form hydronium ion. As a result of this arrangement hydrogen ions coming out of the platinum surface due to the de-electronation reaction will remain close to the surface (on the metal side) without forming a hydronium ion.

8. From quantum mechanical considerations, proton mobility has been explained on the basis of proton tunneling as proton is a fermion and classical theories cannot be applied[21]. It has been further suggested that proton jump can occur only if the acceptor water molecules orient so that the oxygen, the proton and the unshared electron of the acceptor electron line up. The two processes, proton transfer and water re-orientation involve a time of 10^{-14} and 2.4×10^{-13}s respectively. In the present case, viz. hydrogen electrode having platinum as the inert electrode, the exchange reaction is the electronation of hydrogen ion to give hydrogen atom and de-electronation of adsorbed hydrogen atom to give a pro-

ton. Both these reactions can be expected to take place at such times, probably, shorter than 10^{-14} s. Which is the time required for proton transfer. The proton, formed as a result of de-electronation of hydrogen atom on Pt surface, remains without forming hydronium ion due to the reason mentioned above, close to the platinum surface and is readily available for electronation reaction. The electrons formed due to de-electronation of hydrogen atoms absorbed on the metal surface can become members of the conductance of Fermi electrons of the metal. The electronation of the proton (or hydrogen ion) and de-electronation of hydrogen atom is most probably occurring on the electrode surface and close to the inner Helmhotz plane of the electrical double layer populated by highly oriented water molecules. This exchange reaction is almost a solid state reaction and takes place at very high exchange rates that is necessary for a standard reference electrode. Thus the forward and reverse (or backward) reaction, viz. $H^+ + e \rightleftharpoons H$, can take place with zero free energy change i.e. $\Delta G^\circ = 0$ and consequently E° which is given by $-\Delta G^\circ/nF$ is also zero.

The finding provides a solid base for the electrochemical series built up on the intuitive assumption of W. Nernst that the standard potential of hydrogen can be assigned a value of zero.

1. Parthasarathy, S. and Srinivasan, K. R., *Curr. Sci.*, 1994, **66**, 56–59.
2. Sawyer, D. T. and Roberts Jr J. L., in *Experimental Electrochemistry for Chemists*, John Wiley, USA, 1974, p. 36.
3. Jeffery, G. H., Basset, J., Mendeham, J. and Denney, R. C., in *Vogel's Textbook of Quantitative Chemical Analysis*, English Language Book Society, UK, 1989, p. 61.
4. Samuel Glasstone, in *An Introduction to Electrochemistry*, Affiliated East West Press Pvt. Ltd., New Delhi, India, 1975, p. 139.
5. Davies, C. W., *J. Chem. Soc.*, 1938, 2093–2098.
6. Robinson, R. A. and Stokes, R. H., in *Electrolyte Solutions*, Academic Press, New York, 1985, p. 65.
7. Kortum, G., in *Treatise on Electrochemistry*, Elsevier, London 1965, p. 308.
8. Samuel Glasstone, in *An Introduction to Electrochemistry*, Affiliated East West Press Pvt. Ltd., New Delhi, India, 1975, p. 201.
9. Atkins, P. W., in *Physical Chemistry*, English Language Book Society, UK, 1985, p. 354.
10. Maier, C. G., *J. Am. Chem. Soc.*, 1930, **52**, 2159–2170.
11. Maier, C. G. and Raltson, O. C., *J. Am. Chem. Soc.*, 1926, **48**, 364–374.
12. National Bureau of Standards, USAA, 1952, Circular500, p. 1266.
13. Bockris, J. O'M. and Reddy, A. K. N., in *Modern Electrochemistry*, Plenum Press, New York, 1970, vol. 1, p. 461.
14. Bockris, J. O'M. and Reddy, A. K. N., in *Modern Electrochemistry*, Plenum Press, New York, 1970, vol. 1, p. 462.
15. Bockris, J. O'M. and Reddy, A. K. N., in *Modern Electrochemistry*, Plenum Press, New York, 1970, vol. 1, p. 467.
16. Bockris, J. O'M. and Reddy, A. K. N., in *Modern Electrochemistry*, Plenum Press, New York, 1970, vol. 1, p. 465.
17. Antropov, L. I., in *Theoretical Electrochemistry*, Mir Publishers, Moscow, 1972, p. 169.
18. Bockris, J. O'M. and Reddy, A. K. N., in *Modern Electrochemistry*, Plenum Press, New York, 1970, vol. 2, p. 1238.
19. Bockris, J. O'M. and Srinivasan, S., in *Fuel Cells: Their Electrochemistry*, McGraw Hill Book Company, New York, 1969, p. 43.
20. Antropov, L. I., in *Theoretical Electrochemistry*, Mir Publishers, Moscow, 1972, p. 276.
21. Bockris, J. O'M. and Reddy, A. K. N., in *Modern Electrochemistry*, Plenum Press, New York, 1970, vol. 1, p. 487.

Received 8 September 1995, accepted 7 October 1995.

Selective denaturation of cytochrome c oxidase by ionic surfactants: Depletion of both the heme *a* residues from the enzyme

Tapan Kanti Das and Shyamalava Mazumdar

Chemical Physics Group, Tata Institute of Fundamental Research, Homi Bhabha Road, Bombay 400 005, India

UV-visible absorption and circular dichroism studies on interaction of ionic surfactants with bovine heart cytochrome c oxidase have been carried out in lauryl maltoside solution. Results showed that addition of the surfactant causes irreversible depletion of heme *a* from cytochrome *a* as well as cytochrome a_3 sites of this membrane protein at physiological pH. This selective denaturation by ionic surfactants is different from the effect of guanidine hydrochloride on this enzyme. Depletion of heme *a* causes complete inactivation of this ubiquitous respiratory enzyme.

IONIC surfactants, such as sodium dodecyl sulphate (SDS), are commonly used for denaturation of proteins[1–4]. There has, however, yet been no report on any systematic study on the nature of interaction of surfactants with hemeproteins. Moreover, cationic surfactants were considered to be less efficient in such denaturation of proteins, which has been disputed recently[5], lending it subject for further investigation. We have recently shown that surfactants, both cationic and anionic, can cause release of heme-NO species from nitric oxide complex of myoglobin[5]. Previous studies have shown that membrane proteins are generally more resistant to denaturation than other proteins. Cytochrome c oxidase (CcO, EC 1.9.3.1) is a mitochondrial membrane protein[6–11] with two heme *a* active sites, strongly ligated to one or two histidines of the protein. This enzyme is generally extracted in active form in neutral surfactants such as lauryl maltoside (LM), Triton X-100 (TX-100), etc. Structure and function of this multifunctional enzyme is a subject of frontier interest in recent years. The present report shows that the highly buried heme *a* groups of this enzyme can indeed be depleted by treatment with both cationic surfactant, cetyl trimethyl ammonium bromide (CTAB) and anionic surfactant, SDS, in lauryl maltoside solution. The biochemical effect of surfactants is a subject of growing concern and the present study demonstrates their fatal effect on this vital respiratory enzyme. Furthermore, this study provides an

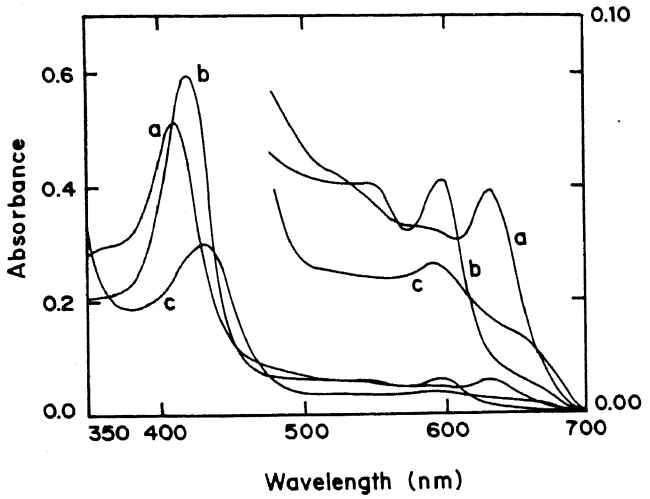


Figure 1. The optical absorption spectra of (a) CTAB-treated CcO. CcO (7.6 μM, 0.1% in lauryl maltoside, 50 mM sodium phosphate buffer, pH 7.4) was charged to a final CTAB concentration of 25 mM. (b) Native CcO spectrum (without CTAB) is given for comparison. (c) Dithionite-reduced CTAB – CcO. Minimal amount of solid dithionite was added to CTAB-treated CcO to effect reduction.

easy method of extracting heme *a* from the enzyme which would otherwise require more drastic conditions. The results have been interpreted in the light of electrostatic and hydrophobic interactions between the surfactant and the enzyme.

Bovine CcO was extracted from ox heart using reported methods[8,12]. The spectrum of bovine heart CcO in lauryl maltoside showed an irreversible change on addition of CTAB or SDS solution to the enzyme. Figure 1 shows the spectrum of CcO (7.6 μM) obtained after addition of 25 mM CTAB (trace *a*) in lauryl maltoside solution. The spectrum of the native enzyme in absence of surfactant is given in trace 'b' (Figure 1) for comparison. The spectrum of CTAB-treated CcO matched exactly with that of CTAB solution of heme *a* separately isolated from the native enzyme (data not shown, see ref. 13). Earlier report[13] of optical spectra of monomeric heme *a* encapsulated in surfactant solution also matched with the spectra shown in trace '*a*' of Figure 1. The Soret peak of CcO undergoes a large blue shift from 421 nm to 410 nm on addition of CTAB. This visible band of CcO at ~ 598 nm also disappears in presence of the surfactant to give a new charge transfer band at 631 nm characteristic of high-spin heme *a* species. Moreover, dithionite reduction of the solution of CTAB-treated CcO (trace 'c' of Figure 1) matched with the reduced heme *a* in CTAB obtained on addition of minimal amount of dithionite to heme *a* solution (data not shown). Similar observations were obtained when CcO was treated with the anionic surfactant, SDS. Unlike ionic surfactants, the neutral surfactants such as, Triton X-100, LM, etc., however, did not have any drastic effect on the spectra of the enzyme. Treatment of guanidine hydrochloride has recently[12] been shown to cause a red shift in the Soret band of CcO. This indicates that the environment around the heme *a* centre on treatment with ionic surfactant is different from that obtained on treatment with guanidine hydrochloride.

The coordination geometry of the two heme *a* centres (cytochrome *a* and cytochrome a_3 sites) in the intact native enzyme are different from each other, giving rise to a slight difference in their optical spectra (broadening of bands) in the native enzyme[9]. However, addition of the ionic surfactant results in complete disappearance of the spectral bands of CcO, indicating that both the cytochrome *a* and cytochrome a_3 sites are depleted from the enzyme to form heme *a* in micelles[14].

In order to further characterize the species obtained on addition of ionic surfactant to CcO, we have measured circular dichroism of the Soret region of CcO at different added concentrations of surfactants. Figure 2 shows that the ellipticity of the Soret region decreases steadily on addition of increasing amounts of CTAB. At surfactant concentrations above 30 mM (with 14 μM CcO) the CD spectra become almost constant with very small ellipticity in the Soret region. The presence of protein environment around the heme gives rise to large ellipticity of Soret band of hemes in hemeproteins[15]. Removal of the heme from the protein and its subsequent encapsulation inside micelles[14] result in an almost isotropic environment around the heme moiety, hence it has negligibly small ellipticity in the Soret region. This result further confirms that the species released on treatment of CcO with the surfactant is heme *a* which gets encapsulated inside the micelles formed by aggregation of the surfactants. The concentration of surfactant required to complete the reaction was found to be almost same for SDS and CTAB, indicating that both of these surfactants are almost equally effective in depletion of heme *a* from CcO. The concentration of guanidine hydrochloride required[12] to diminish the heme CD signal was (> 1 *M*) much more than the surfactant required in the present case, indicating that ionic surfactants are much more effective in releasing heme *a* from the enzyme. Control experiment with pure heme *a*

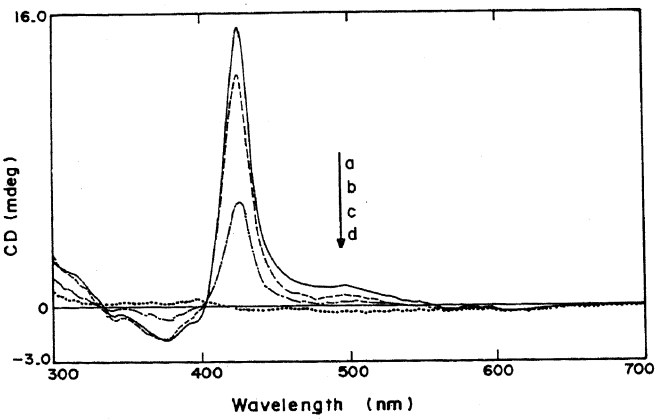


Figure 2. Changes in ellipticity in the visible region of CD spectrum of CcO in 0.1% lauryl maltoside, 50 mM sodium phosphate buffer, pH 7.4, on addition of CTAB aliquots. CcO concentration was 14 μM, cell path length was 10 mm. Curves *a* to *d* represent CTAB concentrations of 0, 12.5, 22.8 and 39 mM respectively.

monodispersed in micelles (data not shown) did not show any significant CD in the near UV-visible region. In the UV region, CD bands around 220 nm were found to diminish up to 36% on addition of the surfactant (data not shown). Similar change in the CD in the 220 nm region has been found on guanidine hydrochloride[12] treatment of CcO. This indicated that the protein possibly gets partially unfolded on treatment with surfactant. However, because of the complexity of the structure of this multisubunit protein, the CD spectra does not provide any quantitative information regarding the unfolding of the structure.

CD and optical spectra of the cyanide complex of CcO (CcO – CN) have also been studied in presence and absence of ionic surfactants. The spectra of the cyanide complex were recorded within one hour after addition of buffered (pH 7.4) KCN solution to CcO when cyanide binds mainly to the cytochrome a_3 site[9]. The major Soret CD peak of CcO – CN was observed at 427 nm, which vanished on gradual addition of SDS or CTAB solution. The optical spectra of surfactant-treated CcO–CN matched quite well with the cyanide complex of heme *a* in CTAB. This result indicates that, addition of ionic surfactants causes breaking of iron–histidine bonds in cytochrome *a* as well as in cytochrome a_3 and results in formation of isolated heme *a* in micellar solution. Walter *et al.*[16] observed a partial spin state conversion of cyt *a* by incubating CcO in higher lauryl maltoside concentration at 40°C. A blue shift of Soret band and formation of a charge transfer band at 630 nm at the expense of the α-band (at 598 nm) was ascribed to a low to high spin transition of cyt *a*, probably because of axial Fe–N bond cleavage[16]. Our results show that this type of transition is complete by treating CcO with SDS or CTAB.

In order to characterize the heme *a* species obtained on treatment in micelles, NMR spectra of heme *a* in CTAB solution at pH~2 (Figure 3, upper spectrum) showed highly down field shifted broad signals due to the heme protons (possibly heme methyl protons). The paramagnetically shifted signals in CTAB solution of heme *a* are similar to those obtained from a $CDCl_3$ solution of heme *a* (lower spectrum in Figure 3), which is typical of high-spin Fe^{3+} heme species. The heme proton signals in micellar solution are generally very broad[14] as observed in the present case. Thus the heme *a* species released on treatment with CTAB on CcO is possibly a high-spin heme *a* complex with axial ligands water and hydroxo ion. The line width of the paramagnetically shifted heme proton signals broadens because of fast ligand exchange at ambient pH, hence they could not be detected at ambient pH using the 500 MHz NMR spectrometer.

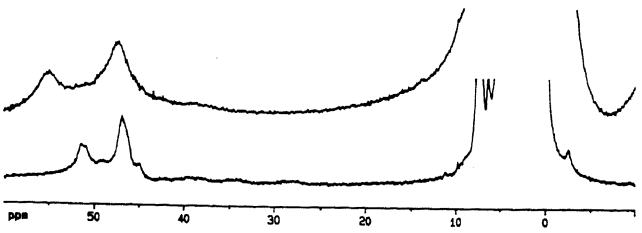


Figure 3. Proton NMR spectrum of heme *a* (~0.4 mM) in $CDCl_3$ (lower trace) and in 30 mM CTAB, 50 mM sodium phosphate buffer, pH~2 (upper trace) in D_2O. Temperature was maintained at 30°C. A multiple frequency irradiation pulse program was used to simultaneously decouple water and CTAB proton resonances.

The present study, thus, demonstrates that ionic surfactants can cause irreversible release of heme *a* from cytochrome c oxidase. The deactivation of CcO and release of heme *a* by CTAB or SDS possibly involves electrostatic binding of the ionic surfactant to the enzyme. Both cationic and anionic surfactants possibly bind to the enzyme surface around its heme crevices to cause heme release. As the charged head groups of surfactant molecules anchor to the enzyme, their hydrophobic tails may interact with the hydrophobic amino acids and break the structure of the enzyme. The iron–histidine bonds possibly become severely weak because of the change in protein structure in the vicinity of heme centres, leading to release of heme *a* out of the enzyme. The heme *a* released from CcO is structurally different from the heme in other hemeproteins. The present method may help in designing heme reconstitution procedures with heme *a* and other hemes in CcO as well as in other hemeproteins.

1. Rao, K. S. and Prakash, V., *J. Biol. Chem.* 1993, **268**, 14769–14775.
2. Singer, S. J. and Nicholson, G. L. *Science*, 1972, **175**, 720–731.
3. Tanford, C. and Reynolds, J. A., *Biochim. Biophys. Acta*, 1976, **457**, 133–170.
4. Helenius, A. and Simons, K., *Biochim. Biophys. Acta*, 1975, **415**, 29–79.
5. Das, T. K., Mazumdar, S. and Mitra, S., *J. Chem. Soc. Chem. Commun.*, 1993, 1447–1448.
6. Saraste, M., *Q. Rev. Biophys.*, 1990, **23**, 331–366.
7. Copeland, R. A. and Chan, S. I., *Annu. Rev. Phys. Chem.*, 1989, **40**, 671–698.
8. Capaldi, R. A., *Annu. Rev. Biochem.*, 1990, **59**, 569–596.
9. Brunori, M., Antonini, G., Malatesta, F., Sarti, P. and Wilson, M. T., in *Heme Proteins* (eds Eichhorn, G. L. and Marzilli, L. G.) Elsevier, NY, 1988, vol. 7, pp.93–153.
10. Gennis, R. B. and Ferguson-Miller, S., *Science*, 1995, **269**, 1063–1064; Tsukihara, T., Aoyama, H., Yamashita, E., Tomizaki, T., Yamaguchi, H., Shinzawa-Itoh, K., Nakashima, R., Yaono, R. and Yoshikawa, S., *Science*, 1995, **269**, 1069–1074.
11. Williams, R. J. P., *Nature*, 1995, **376**, 643; Iwata, S., Ostermeier, C., Ludwig, B. and Michel, H., *Nature*, 1995, **376**, 660–669.
12. Hill, B. C., Cook, K. and Robinson, N. C., *Biochemistry*, 1988, **27**, 4741–4747.
13. Yanagi, N., Sekuzu, I., Orii, Y. and Okunuki, K., *J. Biochem.*, 1972, **71**, 47–56.
14. Mazumdar, S. and Mitra, S., *Struct. Bonding*, 1993, **81**, 115–145.
15. Myer, Y. P., *Curr. Top. Bioenerg.*, 1985, **14**, 149–188.
16. Walter, S. A., Papadopoulos, P. G. and Baker, G. M., *J. Biol. Chem.*, 1993, **268**, 4716–4719.

ACKNOWLEDGEMENTS. We thank Prof. Samaresh Mitra for his constant encouragement and support during the course of this work. NMR work was done at the National Facility of High Field NMR, TIFR, Bombay and we acknowledge help of Ms Uma Joshi and Mr Devidas Jadav in carrying out NMR experiments.

Received 6 July 1995; revised accepted 27 October 1995

Evaluation of haemorheological, biochemical and microcirculatory parameters in young smokers

Ranjeet S. Ajmani*†, R. R. Puniyani and Varjeet S. Khanuja****

*National Institute of Health, Gerontology Research Centre, NIA, LCMB, 4-B-12, 4940 Eastern Avenue, Baltimore, MD 21224, USA
**School of Biomedical Engineering, Indian Institute of Technology, Powai, Bombay 400 076, India

We studied 44 young smokers and 40 controls, all male, for haemorheological, biochemical, physiological and microcirculatory parameters. This study was conducted between January 1994 and December 1994. All the smokers have been smoking 13–18 cigarettes per day for the last three years. None of them were taking any medicine nor did they have any major health problem which could affect the results of our study. The results of our study clearly indicate that smokers have disturbed haemodynamic profile as compared to normal controls. This could result in worsening of blood flow condition, leading to development of various disorders.

THE fact that cigarette smoking affects flow behaviour of blood has been well-documented[1]. The hazard of smoking greatly depends on the way of puffing and inhalation, and also on the number of cigarettes smoked per day[2]. It is composed of more than 4000 substances which have diversified adverse effects on the body. It contains nicotine, tar, polynuclear aromatic hydrocarbons, phenol, cresol and *N*-nitrosonornide, benzol in particulate phase, while the gaseous phase contains carbon monoxide, hydrocyanide, ammonia, acetaldehyde, various oxides of nitrogen and acrolein[3]. Out of these, nicotine is most characteristic of tobacco. It is an alkaloid whose inhalation is known to trigger the secretion of vasopressin, a potent vasoconstrictor hormone, with a possible action on coronary arteries. It also stimulates sympathoneurnal and sympathoadrenal activity and influences systolic blood pressure, heart rate and platelet aggregation. A study conducted by Arnow *et al.*[4] found that smoking high nicotine cigarettes causes rise in blood pressure, heart rate, left ventricular end diastolic pressure, coronary sinus, and in arterial and venous carbon monoxide levels, simultaneously causing a decrease in coronary sinus and in arterial and venous oxygen partial pressure levels with a partial recovery within 30 min. Another component is carbon monoxide, which has an affinity 245 times greater than oxygen and, therefore, reduces oxygen supply to the heart and increases the chances of coronary heart disease. This risk is 20 times if the carboxyhaemoglobin (COHb) level is 5% or more[5]. In patients with underlying coronary heart disease, smoking enhances greatly the incidence of acute myocardial infarction and sudden death. Premature coronary artery disease and mortality arising from it increases progressively with the number of cigarettes smoked and the fatal and nonfatal coronary artery disease (CAD) is 60–70% higher in smokers than in non-smokers.

Due to the more viscous blood, there is a stage of stagnation and deficient oxygen supply to the tissues, leading to hypoxia[6]. This reduced oxygen supply leads to: 1. lowered redox state, 2. rapid glycogen utilization, 3. increased osmolality and swelling of cells, and 4. metabolic acidosis.

Metabolic acidosis has a very strong negative effect on the structure and functions of red blood cells. The slower the blood flow in the capillaries, greater is the internal viscosity of the red blood cells and greater the chances of breakdown of haemodynamic mechanisms at macro-and microcirculatory levels. This decreases further the oxygen supply to the tissues, leading to hypoxia and further decrease in pH. Hypoxia stimulates the production of red blood cells, i.e. increase in haematocrit. Haematocrit is one of the strong determinants of blood viscosity. Gagnon *et al.*[7] found that increase in haematocrit coupled with blood viscosity has serious effects or organ perfusion. Thus, a vicious circle sets in which leads to further deterioration of blood flow, making smokers more susceptible to development of vascular disorders.

Indian Institute of Technology (IIT) Bombay is a residential campus where graduate and postgraduate students stay in different hostels. The total student population is about 2500. For conducting this study, students from different hostels were contacted and briefed about the nature of the study. About 152 students expressed their willingness to participate in this study. Only 44 smokers who did not have any major health problems and at the same time did not have any other addiction like alcohol/tobacco chewing, etc., were selected for this study. None of them were taking any treatment, particularly aspirin, at least two weeks before the laboratory tests were performed. All of them were smoking 13–18 filtered cigarettes/day for a variable period of 2–3 years. Out of 152 students, 40 were selected as control group, as they had not smoked cigarettes so far and also were clinically normal subjects. For getting their details and thorough medical, family and personal histories, a questionnaire was used. On the day of blood collection, their blood pressure, pulse rate and weight was taken. Arm blood pressure was taken under supine condition after 10 min of rest, using sphygmomanometer and peripheral

†For correspondence

Table 1. Details of the subjects

Group	Number	Age	Weight	Pulse	(SBP/DBP mm Hg)
NC	40	23.06 ± 1.86	65.23 ± 8.22	77.45 ± 9.01	130.6 ± 8.33 84.7 ± 3.9
SM	44	22.95 ± 2.14	67.93 ± 10.76	88.56 ± 10.19	145.2 ± 5.19 87.3 ± 5.83

NC: Normal controls;
SM: Smokers
SBP/DBP: Systolic blood pressure/Diastolic blood pressure;
Subjects from both categories were males.

pulse was palpated. To minimize variability, one subject from each group was examined each morning. Details of these subjects are given in Table 1.

Blood collection: About 10 ml of venous blood was drawn from antecubital vein using a plastic syringe with 21-gauge stainless steel needle with minimal suction. Out of 10 ml, about 6 ml was anticoagulated with dipotassium salt of ethylene diamine tetraacetic acid (EDTA) 1.8 mg/ml and about 4 ml with citrate (3.8%). Citrated blood was used for platelet aggregation study. The blood sample of the subjects was collected in the morning between 9.30 and 10 a.m. (to avoid any diurnal variations), after 8 h of fasting, in the Laboratory of Haemorheology and Microcirculatory Sciences, School of Biomedical Engineering, IIT, Bombay.

The sample was analysed within 2 h of collection as per recommendation of International Committee for Standardization in Haematology[8–10]. The following parameters were investigated:

Hemorheological parameters: Whole-blood viscosity, plasma viscosity, red-cell rigidity, red-cell aggregation and platelet aggregation.

Biochemical parameters: Haematocrit, erythrocyte sedimentation rate, total proteins, fibrinogen, albumin, cholesterol and triglycerides.

Physiological parameters: Blood pressure, pulse and weight.

Microcirculatory parameters: Transcutaneous oxygen tension and skin blood perfusion.

Whole-blood viscosity, plasma viscosity and red cell rigidity were measured by using Contraves LS-30, as performed by Ajmani *et al.*[11]. For the measurement of red cell aggregation, haematocrit, platelet aggregation, erythrocyte sedimentation rate, cholesterol triglycerides and plasma protein the method of Puniyani *et al.*[12] was used. Biochemical parameters like plasma albumin, total proteins, cholesterol and triglycerides were estimated by using M-10 autoanalyser. Fibrinogen estimation was done using the method of Ratnoff and Manzie[13], where the fibrinogen present in the plasma is converted to fibrin by the action of thrombin and subsequently boiled with sodium hydroxide. Its tyrosine content is estimated using the Foil–Ciocalteu reagent.

For simultaneous flow dynamics and oxygenation studies, the subjects were called in the evening as they had to attend the regular course classes in the morning. Moreover, since the study of these parameters takes at least 2 h for each subject, evening was the most convenient time to carry out this study. These studies were carried out in nonfasting state.

The Periflus PF 3 laser Doppler perfusion monitor (LDPM) was used for continuous monitoring of local microvascular blood flow as per Schubert *et al.*[14].

The transparent transcutaneous triple sensor/monitor was used for measurement of oxygen tension ($TcPO_2$) and, simultaneously, flux pattern (flow/velocity) by laser Doppler equipment. The TTC-45 triple sensor is compatible with Periflux PF-3. Oxygen tension was measured by AVL triple sensor/monitor TTC-45 at 44°C on three anatomical sites (right ankle, right wrist and right elbow). AVL triple sensor TTC-45 takes 10–15 min to stabilize at each site and hence the readings are taken exactly after 15 min, when the slope of the curve becomes constant.

The $TcPO_2$ measurement is basically a polarographic determination, by means of a modified Clark-type electrode, of molecular oxygen diffusing into the skin surface. Local hyperaemia is necessary to measure PO_2 on the skin surface, which has a close relationship with change in arterial PO_2. The hyperaemic condition is achieved by heating the silver/silver chloride anode to 45°C. The PO_2 signal results from the reduction current of the platinum cathode, which is proportional to the amount of oxygen diffusing into the skin surface.

The results for smokers were compared with normal controls as nearly as possible for age and sex. Student's *t* test was used for the analysis of the data. Multivariate correlation analysis was carried out by standard statistical techniques.

The results of this study are presented in Table 2 and 3.

1. There was an increase in the whole-blood viscosity at both high and low shear rates. But changes were more significant at high shear rates ($p < 0.001$ to 0.004).
2. Increase in plasma viscosity was very significant ($p < 0.001$).
3. Red cell rigidity parameter was most significantly ($p < 0.0001$) affected as compared to other parameters.
4. Fibrinogen had the highest statistical significance among the biochemical parameters ($p < 0.001$).

Table 2. Means and standard error of haemorheological and biochemical parameters

Parameters	A-group (40)	B-group (44)
Whole blood viscosity (cP)at shear rate (s^{-1})		
94.5	4.32 ± 0.79	5.63 ± 1.03#
51.2	4.62 ± 1.03	6.44 ± 1.52#
20.4	5.44 ± 1.76	7.93 ± 2.01#
8.11	7.05 ± 1.23	10.53 ± 1.93**
4.39	8.76 ± 2.32	13.37 ± 3.56##
3.32	10.11 ± 2.15	14.94 ± 3.42**
1.28	15.74 ± 3.45	23.36 ± 4.36**
0.51	28.05 ± 4.34	37.52 ± 5.76**
Plasma viscosity (cP) at 51.2/s	1.29 ± 0.04	1.59 ± 0.19**
Red cell rigidity (cP) at 94.5/s	3.53 ± 0.61	4.76 ± 0.63***
Red cell aggregation	12.34 ± 3.76	17.22 ± 3.21**
Haematocrit (%)	43.29 ± 3.14	55.11 ± 5.48###
Fibrinogen (mg%)	311.33 ± 20.87	387.6 ± 21.74**
Erythocyte sedimentation rate (mm/h)	40.55 ± 8.22	90.34 ± 15.43###
Triglycerides (mg%)	88.43 ± 11.13	125.65 ± 10.78$
Cholosterol (mg%)	159.33 ± 15.21	190.28 ± 17.90$
Albumin (g%)	3.66 ± 0.11	4.22 ± 0.47
Total proteins (g%))	7.94 ± 0.78	8.36 ± 0.61

***$p < 0.0001$, **$p < 0.001$, *$p < 0.01$, ##$p < 0.004$. #$p < 0.003$, ###$p < 0.005$, $$p < 0.05$.
A-group: Normal controls.
B-group: Smokers.

Table 3. Microcirculatory parameters

Parameters	Non-smokers	Smokers	P
$TcPO_2$(1)	75.65 ± 11.01	62.90 ± 9.01	< 0.05
$TcPO_2$(2)	84.57 ± 15.20	68.75 ± 11.28	< 0.05
$TcPO_2$(3)	76.36 ± 7.98	65.46 ± 11.44	< 0.05
BL.PR. (1)	163.07 ± 38.77	124.79 ± 54.78	n.s.
BL. PR. (2)	257.26 ± 77.38	201.04 ± 73.32	n.s.
BL. PR. (3)	203.09 ± 62.02	191.67 ± 62.72	n.s.

$TcPO_2$: Transcutaneous oxygen tension;
BL.PR.: Blood perfusion;
n.s.: non-significant.
(1) At ankle; (2) At wrist; (3) At elbow.

5. Red cell aggregation had a significance of $p < 0.001$.
6. Erythrocyte sedimentation rate and platelet aggregation were also elevated ($p < 0.005$).
7. Haematocrit, triglycerides and cholesterol levels were also raised, with significance values of $p < 0.005$ and $p < 0.05$, respectively.
8. Total proteins and albumin fraction were unaltered.
9. Pulse rate was higher than normal controls. In a few cases the pulse was irregular and blood pressure readings, particularly diastolic, were on the higher side of the normal.
10. Skin blood perfusion and transcutaneous oxygen tension at all the three anatomical sites were found to be low. Although skin perfusion was reduced, it was not statistically significant.

Belch *et al.*[15] have shown that smokers have elevated levels of fibrinogen ($p < 0.04$), lower plasminogen ($p < 0.02$) and higher plasma viscosity ($p < 0.003$). Just after smoking three cigarettes there was a substantial increase in platelet aggregation ($p < 0.02$), alpha2 M ($p < 0.02$) and factor VIII RAG ($p < 0.05$). A prospective study conducted by Galea and Davidson[16] and Rampling *et al.*[17] on smokers reported significant improvement in overall haemorheological and haematological variables after 2 weeks of discontinuation of smoking. Billimoria *et al.*[18] divided smokers into two subgroups based on the number of cigarettes being smoked per day (> 20 cigarettes/day: heavy smokers, and 5–15 cigarettes/day: light smokers). Heavy male smokers showed higher fasting serum turbidity, cholesterol, phospholipid, triglycerides, esterified fatty acid index of beta and pre-beta lipoprotein and reduction in clotting time compared to normal controls. In the case of the women's group, though there was increase in triglycerides, pre-beta esterified fatty acid index, longer fibrinolysis time and reduction in clotting time, no change was observed in cholesterol, beta esterified fatty acid index and fasting serum turbidity. In the case of male heavy smokers, haematocrit, haemoglobin and mean corpuscular

volume were elevated. The white cell count was very significantly raised; in particular, there was a rise in neutrophils and lymphocytes. But these changes were not observed in the women's group. In the case of light smokers, changes did not achieve statistical significance but the trend was similar. We found increase in the blood viscosity at all the shear rates considered, although changes were more significant at high shear rates than at low shear rates. This indicates prime importance of red cell rigidity among smokers. Parallel to this, rise is haematocrit also contributes to rise in blood viscosity. There was significant positive correlation (0.846) between blood viscosity and haematocrit, indicating its role in haemoconcentration. Rise in haematocrit could be a response to consistent hypoxic state among smokers. Changes at low shear rates were mainly due to increase in red cell aggregation. Increase in the red cell aggregation was mainly due to rise in fibrinogen level and plasma viscosity. Increase in plasma viscosity is attributed to rise in whole-blood viscosity as there was positive correlation between them at all the shear rates. In our study we found rise in platelet aggregation over the range 48–76%. The rise in erythrocyte sedimentation rate among smokers could be due to the presence of many charged particles released in the blood while one is smoking[19].

Among the biochemical parameters, change in triglycerides and cholesterol was significant. The increase in cholesterol has an adverse effect on red cell rheology as well as plasma viscosity. These changes were coupled with changes in microcirculatory parameters, reduction in peripheral skin blood perfusion and transcutaneous oxygen tension. Although it is premature to comment on whether these microcirculatory changes precede or follow the haemorheological changes, if we consider this in totality, it is apparent that smokers have disturbed microcirculatory flow. Reduced skin blood supply along with low oxygen partial pressure will hamper perfusion processes, resulting in accumulation of metabolites, leading to haemorheological changes. Norton and Rand[20] have reported a reduction in the surface to volume ratio in red blood cells among smokers. It is indicative of increased red cell rigidity. However, Baker *et al.*[21] found no change in the red cell deformability among smokers. Hawkins[22] in his study found that platelet aggregation can be induced at lower concentrations in smokers than in nonsmokers. The coagulation time of whole blood in smokers were significantly shorter at 0.001% level. The thromboelastograph also reflects similar results. No difference in platelet counts is indicative of qualitative changes in the functioning of platelet cells in this context. In smokers the increase in maximum tensile strength of a clot was observed by 5% compared to nonsmokers. This indicates that there is a greater risk of formation of haemostatic plugs. This could offer a probable explanation of why smokers have higher rate of cardiovascular morbidity and mortality.

The microcirculatory changes mediated through vasopressin were studied by Waeber *et al.*[24]. In their study on 20 male volunteers using laser Doppler flowmetry, reduction in blood flow was observed by induction of concomitant stimulation of vasopressin secretion.

In conclusion, blood viscosity and microcirculatory parameters should be included in routine medical check-up for early detection of silent clinical conditions associated with smoking.

1. Yong, Y. D., in *Horizon in Haemorheology* - 1, Shanghai Scientific and Technological Literature Publishing House, Shanghai, 1990, pp. 145–155.
2. Helmut, L. and Ehrly, A. M., *Clin. Hemorheol.*, 1981, **3**, 241–249.
3. Kar, C. C. and Das, M. K., *Indian Heart J.*, 1989, **41** (2), 72–74.
4. Arnow, W. S., Cassidy, J., Vangrow, J. S., *et al.*, *Circulation*, 1974, **50**, 340–347.
5. Report Surgeon General: Smoking and health - Washington DC., Deptt. of Health Education and Welfare, Public Health Services Publication No. 79-500066, 1979.
6. Wasserman, K., *Circulation*, 1989, **80**, 1084–1086.
7. Gagnon, D. R., Zhang, T. J., Brand, F. N. *et al. Am. Heart J.*, 1994, **127**, 674–682.
8. International Committee for Standardization in Haematology Expert Panel on Blood Rheology, *Clin Hemorheol.*, 1986, **6**, 339–353.
9. International Committee for Standardization in Haematology Expert Panel on Blood Rheology, *J. Clin. Pathol.*, 1988, **41**, 1203–1212.
10. International Committee for Standardization in Haematology, *J. Clin. Pathol.*, 1993, **46**, 198–203.
11. Ajmani, R. S. and Puniyani, R. R., *Clin. Hemorheol.*, 1995, **15**, 347–357.
12. Puniyani, R. R. and Sameer, Sonar, *Clin. Hemorheol.*, 1991, **11**. 397–403.
13. Dacie, J. V. and Lewis, S. M., in *Practical Haematology*, Churchill Livingstone, Longman, XI edn, 1985, p. 229.
14. Schubert, V. and Fagrell, B., *Clin. Physiol.*, 1989, **9**, 535–545.
15. Belch, J. J. F., Ardle, B. M., Burns, P, Lowe, G. D. O. and Forbes, C. D., *Thrombo. Haemosta.*, 1984, **5li**, 6–8.
16. Galea, G. and Davidson, R. J. L. *J. Clin Pathol.*, 1985, **38**, 978–984.
17. Rampling, M. W., Brown, J. R., Robinson, R. J. *et al.*, *Clin Hemorheol.*, 1991, **11**, 441–446.
18. Billimoria, J. D., Pozner, H., Metselaar, B. *et al.*, *Atherosclerosis*, 1975, **21**, 61–76
19. Donner, M., Muller, S. and Stoltz, J. F., *Clin. Hemorheol.*, 1982, **2**, 109–115.
20. Norton, J. M. and Rand, P. W., *Blood*, 1981, **5**, 671–675.
21. Baker, C. W., Anderson, K. W., Beaty, M. M. *et al.*, *Clin. Hemorheol.*, 1993, **13**, 671–678.
22. Hawkins, R. I., *Nature*, 1972, **236**, 109–111.
23. Robinson, K, Conroy, R. and Mulcany, R., *Br. Heart J.*, 1988, **60**, 465–469.
24. Waeber, B., Schallear, M. D., Nussberger, J. *et al.*, *Am. J. Physiol. Heart Circul. Physiol.*, 1984, **247**, H895–H901.

ACKNOWLEDGEMENTS. We thank IIT authorities for providing funds and necessary support from IRCC budget to carry out this research work. We also thank Mrs R. V. Gupte, Mrs S. Patankar and Mrs D. Doshi for their active help in carrying out the investigations for this study. We are also thankfull to all our young friends who have extended their helping hand for this research work and agreed to undergo this study. The work was carried out at the School of Biomedical Engineering, IIT (Bombay), India.

Received 29 April 1995; revised accepted 2 September 1995.

Law, Ethics and HIV, Proceedings of the UNDP Intercountry Consultation. Robert A. Glick, ed. UNDP, New Delhi. 246, pp., 1993.

AIDS is big news today, disturbing and fearsome to people everywhere. The fact that most people are unaware of the potential threat for which there is no vaccine nor cure makes the problem challenging to policy planners and decision makers in the areas of health, human rights and social security. Because of the prospect of undermining social and economic development, HIV/AIDS epidemic is perceived as a development problem to be attached through available strategies of prevention and containment. Little is achieved by generating a scare in society, though the problem warrants nothing less from all concerned. Still less is accomplished by sitting back and lamenting on the catastrophe about to befall humanity, particularly in those parts of the world where poverty and backwardness have already put life in jeopardy for many people. Statesmen are worried about nuclear weapons, and their proliferation in the world. They, therefore, put nuclear non-proliferation issue as the number one item in the international agenda. They are probably unaware of the silent but sure killer developing in their backyard for which no treaty nor declaration can be an adequate response. The HIV/AIDS issue is at once a challenge and an opportunity for united action for the entire world to prevent human suffering and promote the ultimate human right to life and happiness.

If the problem is overstated, it is deliberately so done. Even educated people refuse to believe the terrible consequences of this dreaded disease. When the potential threat to human survival from the HIV/AIDS was narrated in a social gathering recently, an old lady in the neighbourhood paused and added that it might well be so because such a calamity is predicted in the scriptures several centuries ago. A feeling of helplessness and a sense of inevitability seem to be the reaction of the old and elderly. On the other hand, the younger generation refuse to believe it to be so serious and think that modern medicine does have a cure for every illness, including AIDS. The medical profession, at least in India, seems to have not taken up the issue with the attention it deserves and the media is indifferent excepting on occasions when a sensational story makes news. In this context, the initiative of the Regional Bureau for Asia and the Pacific of the United Nations Development Programme to study the legal, ethical and policy framework in different societies in the region is to be welcomed. Information is the key for intelligent policy-making. Law can succeed only when it is based on correct information as to the behavioural patterns, motivations and beliefs of people whose conduct it is supposed to influence. It is, therefore, of great advantage to the countries in the region that the volume under review has brought together a wealth of information which provides at least a base for further inquiry and research to guide policy on this vital area of public health and human rights.

The volume essentially is a collection of papers written by experts organized around issues and countries in the Asia–Pacific region. A third section addresses questions on strategies and priorities, giving policy choices in the available legal framework. The foreword by Justice Michael Kirby comprehends the objective and thrust of the whole exercise. A collective appreciation of the strengths and weaknesses of the myriad legal regimes of the region in responding to the unlimited peril of HIV/AIDS resulted in the identification of a variety of strategies varying from increased availability of condoms, safe sexual practices through family education and equal states of women in decision-making processes. More popular education and better policies on drug use, public health and prostitution are recommended by the paper writers as possible initiatives in preventive action. The object of networking in the region for dissemination of information on policies and programmes directed towards HIV/AIDS is indeed a commendable outcome insofar as not much is known about the disease and about how to counteract its spread, given the attitudes and behaviour towards sex of the people. As the editor of the volume expects, the papers should assist in defining a research agenda for the future and in assembling a practical action programme.

For a student of law, the HIV/AIDS debate raises a number of issues for interrogation. It is accepted that, in the interest of public health, individual rights can be restricted. In the present context, can we articulate the limits and limitations which can possibly be imposed on freedom and liberty? How do religious freedom and minority rights in a democratic, pluralist society get affected by public health considerations and to what extent are they constitutionally justified?

Regulating sex behaviour has been a most difficult and controversial area of law. It is common knowledge that observance of safe sexual practice is the best strategy for prevention of HIV/AIDS. To what extent can law facilitate such practices? Under what circumstances can law punish unsafe sexual behaviour? Should prostitution be prohibited, regulated or legitimized in the interest of safe sex? Can HIV-infected persons be isolated and their sexual activity controlled?

Issues are complicated in several jurisdictions in the region because of the prevalence of customary, cultural and religious practices not always conducive to safe sex approaches. Ethical and moral standards of communities are developed not necessarily on rational and scientific arguments. They are of mixed origin and followed by habit without challenge. Obsensantism, superstition and exploitative customs complicate the situation, posing serious challenges to social engineering through law.

In an incisive essay contributed by Julie Hamblin on the role of law in developing HIV/AIDS policy, three models are suggested, namely, the proscriptive model, where conduct is influenced by penal sanctions, the protective model, wherein the law tries to uphold the rights and interests of those affected or likely to be affected by HIV/AIDS through rights-sensitive legal responses; and the instrumental model, where law is a proactive instrument to advance social policies through value changes. Keeping in mind the record of criminal law, it is argued that the proscriptive model could become counterproductive. Compulsory testing and reporting with concomitant sanctions for noncompliance may impede preventive efforts. Nonetheless, the temptation for policy makers is to go for the quick and easy response of prohibition and penalization. Today, when decriminalization in drug use and in prostitution is widely canvassed, the proscriptive role of law has to be invoked

for negative prohibition rather than for positive criminalization.

Imaginatively used, law has a significant role in promoting the desirable conduct by intelligent arrangements of incentives and disincentives in a broad spectrum of activities. There are many examples of such use of law in society and the scope of similar use in the control of HIV/AIDS needs to be examined. The proactive role of law is one which requires a variety of informational input, policy coordination, monitoring, research and development. It envisages a sophisticated legal system with established institutions and standards. A law reform process should be built into the system. Interaction among different professionals, continuing education for actors in the system and adequate research and media support are necessary for law to influence social change in complex issues of human behaviour. Every society does use law in this role with varying degrees of success. Here law can well reinforce other forces of change in society. According to Julie Hamblin, the instrumental role of law to buttress other HIV/AIDS strategies is critical in the 1990s.

The papers raising these issues make interesting reading, particularly for those from outside the countries of their reference. A case study on labour legislations and labour relations from the Phillippines seeks more specific standards and guidelines from the law to compel employers to adopt more enlightened and human-rights-friendly approaches. Another case study from the same country examines the law on privacy and information in the context in HIV/AIDS and projects its potential for influencing positive behavioural change.

Family law buttressed by religious beliefs and the unequal and discriminatory status of women, particularly in family relationships, are the issues examined by participants from Malaysia and India, respectively. They bring out a variety of inadequacies in the existing law which inhibit proper awareness and positive responses to the AIDS syndrome.

The section on country papers includes ten contributions from as many countries in the region. They are a valuable source of information on the existing legal framework and to plan further action in the legal and ethical fronts. Comparative experiences from different countries are of great value to researchers and policy planners, particularly when the problem affects the entire humanity across political boundaries.

The third section, which claims to contain policy/strategy papers, is perhaps the weakest segment of the publication. Much of the ideas contained are repetitive and placement-rhetorical. The presentation of HIV epidemic as a development issue is not convincing enough. Critiquing the health care systems in the context of the demands of individual rights does not make much sense in the framework of a market-driven economy. However, a beginning is made by articulating the problem and questioning some of the established public health policies. Law has to be used, but with caution and restraint.

On the whole, UNDP Regional Project has done a great service in bringing the issues together and inviting the attention of researchers and policy makers, health administrators and human rights activists to the implications of an epidemic which may be the worst human disaster unless preventive strategies are adopted well in time. This is an issue as important as nuclear disarmament or environment protection. It deserves as much attention to world leaders and scientists as is being given to environment and nuclear arms. Perhaps the Secretary General may consider declaring the UN year for AIDS Awareness and Prevention and may take steps to convene a World Summit before the turn of the century to evolve concerted action to face the challenge.

N. R. MADHAVA MENON

National Law School of India University, Nagarbhavi, Post Bag No. 7201, Bangalore 560 072, India

The Collected Papers of Albert Einstein - Volume 5 - The Swiss Years: Correspondence, 1902–1914. English Translation, Princeton University Press, Princeton, New Jersey. 1995.

Well before his passing away in 1955 at the age of seventy-six, Albert Einstein's name had become a household word throughout the world. Aptly characterized as 'creator and rebel', through his work he profoundly altered forever our understanding of the natures of space and time; created a new theory of gravitation unparalleled for its beauty and significant for bringing geometry once more into the domain of science; and with Planck before him and Bohr after, opened the doors to the stunningly unfamiliar and rich quantum view of nature.

Already in 1950 Einstein had signed and sealed his last will and testament. His secretary since 1928, Helen Dukas, and his friend Otto Nathan were named joint trustees of his estate, with the latter being sole executor. Possibly due to their attitudes and perceptions of their responsibilities, the world has had to wait for almost four decades before Einstein's papers became available for open study and publication. Predictably, in the last couple of years and in tune with today's 'bare all' journalistic tendencies, there has appeared a rash of accounts 'exposing' Einstein's affairs, especially his relation to his first wife Mileva. On the scientific side, his collected works are yet to appear in final and definitive form; in contrast, those of Neils Bohr have been available now for a quarter century or so.

The book under review is the fifth volume of Einstein's collected papers, being published in English translation by Princeton University Press. It covers the Swiss years, 1902 to 1914, and is a collection of over five hundred letters written to or by Einstein. At about the beginning of this period, in early 1901, he had become a Swiss citizen. He came to Bern in February 1902 and a few months later, as all the world knows, he obtained employment in the Patent Office as 'Technical Expert, Third Class'. His father died later that year, and he married Mileva in January 1903. His sons Hans Albert and Eduard were born in 1904 and 1910 respectively. After seven years at the Patent Office, in 1909 he resigned to become Associate Professor at the University of Zurich. Soon there was a brief move to Prague, then a return to Zurich but this time as Professor at the ETH. This too was not to last very long, as an irresistible offer took him away to Berlin in early 1914. Also, towards the end of this period, the relationship with Mileva was breaking up, and he was getting very

close to his cousin Elsa Lowenthal, whom he was to marry in 1919.

In his physics, we see Einstein at the start of his career, with rare independence and boldness in his conceptions, slowly gaining recognition from his peers, until he matured into the acknowledged and unrivalled master. The period of this volume covers 1905—the 'annus mirabilis' when he sent his three epoch-making papers on Brownian motion, special relativity and the photon concept to the *Annalen der Physik*; the holding in 1911 of the first of many memorable Solvay Congresses in Brussels; and the initial insights and understanding of relativistic gravitation which culminated in 1916 in his general relativity theory.

Einstein's major correspondents in these five hundred and odd letters were his friends from youth, a few members of the family and some of the leading physicists of those times including his seniors. Such a volume has certain inevitable drawbacks as one cannot read the contents as one might read a novel or biography or a properly written account of some scientific topic. There are gaps, often an answer to a letter no longer available, so one is left guessing. Nonetheless one gains a general impression of the growth of Einstein's scientific ideas, the development of personal relationships, the steady increase in his confidence and depth of understanding and independence and strength. We also see how scientists interacted with each other in those times. We are reminded too that Einstein was not always a grey eminence with a saintly halo, but was once a young man on the threshold of life, with all the feelings natural to such an age.

Given the scale of his accomplishments, Einstein is often compared to Isaac Newton. What emerges from these letters is that Einstein had really a very easy-going nature, willing to joke with others in a light-hearted and playful manner, and no signs of malice towards anyone—all in all, a contrast to the somewhat dark, secretive and vindictive Newton. He seems to have had a balanced view of people and circumstances, and an ability in many ways to distance himself and stand apart in judging any situation. Also of course the strength to get to the heart of any scientific matter and to be concerned with the fundamentals.

On the family side, one is surprised to see not a single reference in any letter to the death of Einstein's father in October 1902. Indeed, even letters between Albert and his mother Pauline are quite rare, and he was very much aware of the strains between her and his wife Mileva. There is a 1903 letter of Einstein to Mileva, which refers to a daughter Lieserl born to them before marriage. In the early exchanges between them, one sees much affection and use of pet names, much warmth. But as the years go by, these wane. Towards the end of this volume there are many letters from Einstein to Elsa, though few replies, bringing out his dependence on and need for her company. A quite amusing letter to one Georg Meyer asks the latter to 'bear no grudge against your good wife, who did nothing wrong—a reference to an early attachment of Einstein's? All this goes to reinforce the contrast with Neils Bohr's family background—a stable and caring academic atmosphere, and a deep and cordial marriage that lasted many decades.

On the scientific side, we see Einstein's early concerns with the foundations of thermodynamics, attempting to develop them from the atomic-molecular point of view. These are generally not as well known as Boltzmann's and Gibbs' endeavours; they remind us sharply that there was a time when the reality of atoms was suspect, and that too not so long ago. We also see how he was simultaneously grappling with diverse foundational problems—thermodynamics, electrodynamics, nature of space–time, nature of radiation, the quantum puzzle—none of which would 'leave him alone'. It was a ceaseless struggle to penetrate the secrets of nature, on many fronts, not always only the deep and beautiful but also often the most practical. His interests in problems of chemistry, bringing out again the molecular basis of matter, were recurring. As a young man, his letters to both Lorentz and Sommerfeld show the greatest of respect. Indeed we can say he regarded the former as a kind of father-figure. Describing Lorentz's handling of the sessions at the first Solvay Congress, Einstein says 'Lorentz chaired the conference with incomparable tact and unbelievable virtuosity'; and elsewhere calls him 'a work of art'. In an early letter to Sommerfeld, in 1908, Einstein expresses the wish that he would like to 'perfect my knowledge of mathematical physics' by sitting in on Sommerfeld's lectures. The early letters to Stark and Lenard are also very respectful; but thereafter references to the latter are quite disparaging. These two were later to be in the forefront of the antisemitic attacks on Einstein.

Many of the letters are detailed exchanges of views on scientific problems in a leisurely and open manner that in today's competitive atmosphere are practically unthinkable. The struggle with the photon idea, and Lorentz's detailed analysis and criticism of Einstein's conceptions, are most revealing, though with today's understanding they may seem not very relevant. One misses however, and sadly, the three 1905 letters of submission to the *Annalen der Physik*. On the professional side, one is treated to the hilarious game of musical chairs among the German-speaking universities of the day—a new occupant of a chair at one place immediately waiting to 'receive a call' to move to a better place. During Einstein's brief tenure at Prague, he expressed himself quite frankly against the bureaucracy, stifling traditions and poor quality of students of the place. The several efforts to lure Einstein to Holland, at times involving Lorentz, are part of this drama of constant movement.

Towards the end of this volume comes the phase when Einstein was working feverishly, sided by Marcel Grossmann, to create his new theory of gravitation—the many insights, false leads, the realization that mathematical structure has a great role to play alongside physical reasoning. And on the personal side, the growing relationship with Elsa. To the reader too, hopefully, should come the realization that these letters, and especially the earlier ones which were presumably written with no inkling that they would one day achieve archival value, ought to be read with consideration and a sense of humility.

N. MUKUNDA

Centre for Theoretical Studies and Department of Physics, Indian Institute of Science, Bangalore 560 012, India

INDIAN INSTITUTE OF TROPICAL METEOROLOGY

DR. HOMI BHABHA ROAD, PASHAN, PUNE 411 008

(An Autonomous Research Organization under the Department of Science and Technology, New Delhi)

SPECIAL DRIVE FOR RECRUITMENT OF SC/ST CANDIDATES

Applications in the following format are invited from SC/ST candidates only for recruitment to the following reserved posts:

1. Name in full (in block letters); 2. Mailing address; 3. Sex; 4. Date of birth and age as on 1. 1. 1995 (documentary evidence to be enclosed); 5. Father's name; 6. Whether belongs to SC/ST (documentary evidence to be enclosed); 7. Nationality; 8. Educational qualifications (in chronological order from matriculation onwards) (enclose a separate sheet, duly authenticated by your signatures) documentary evidence is to be enclosed; 9. Experience (if any) with field of specialization 10. Employment record (separate sheet to be enclosed); 11. Whether a Govt. servant?; 12. Give names and addresses of two referees not related to you; 13. List of enclosures.

1. SENIOR SCIENTIFIC OFFICER GRADE-II :

Three posts (one post reserved for SC and two posts reserved for ST candidates)

Scale of pay	:	Rs. 2200–75–2800–EB–100–4000.
Age limit	:	35 years or below as on 1. 1. 1995
Qualifications	:	*Essential*: (i) A good Master's degree or equivalent in Meteorology, Physics, Mathematics, Applied Mathematics, Statistics, Applied Physics, Geophysics or Electronics or BE/B Tech/M Tech (Electronics/ Electrical Engineering) degree. (ii) About 3 years research experience (as evidenced by published papers). *Desirable*: Knowledge of numerical techniques as applied to atmospheric numerical models/statistical methods and computer programming/data base management/systems programming.

2. JUNIOR SCIENTIFIC OFFICER :

One post (Reserved for ST candidate)

Scale of pay	:	Rs 2000–60–2300–EB–75–3200–100–3500
Age limit	:	30 years or below as on 1. 1. 1995
Qualifications	:	*Essential*: A good master's degree or equivalent in Meteorology, Physics, Mathematics, Applied mathematics, Statistics, Applied physics, Geophysics or Electronics or BE/ B Tech./ M Tech. (Electronics/ Electrical Engineering) degree. *Desirable*: Knowledge of numerical techniques applied to atmospheric numerical models/Statistical

methods and computer programming/data base management/systems programming.

3. SECTION OFFICER :

One post (Reserved for ST candidate)

Scale of pay : Rs 2000–60–2300–EB–75–3200–100–3500

Age limit : 35 years or below as on 1. 1. 1995

Qualifications : *Essential*: (i) Bachelor's degree, (ii) About 5 years experience in Accounts/Administration/Purchase and Stores work in a Scientific/Technical/Govt./ Semi-Govt. organization of which at least 3 years should be in a supervisory grade.

4. JUNIOR TECHNICAL ASSISTANT (Computer)

One post (Reserved for ST candidate)

Scale of pay : Rs 1320–30–1560–EB–40–2040

Age limit : Minimum 20 years and maximum 25 years as on 1. 1. 1995

Qualifications : *Essential*: A degree in Science or a diploma in Engineering or a National Apprenticeship certificate or an ITI Certificate (after matriculation) or a matriculation with minimum 3 years experience in Computer Programming.

Desirable: Knowledge of PC, software packages like DBASE, WORDSTAR and FORTRAN Computer programming.

5. STENOGRAPHER GRADE-III :

One post (Reserved for ST candidate)

Scale of pay : Rs 1200–30–1560–EB–40–2040

Age limit : 18–25 years as on 1. 1. 1995

Qualifications : *Essential*: SSC or equivalent with minimum speed of 40 wpm in English typing and 100 wpm in English shorthand

Desirable: Good command over English and knowledge of typing on Wordstar.

The posts carry other allowances admissible to Central Govt employees stationed at Pune.

Upper age limit is relaxable by 5 years for candidates as per rules for ex-servicemen and physically handicapped persons. No age limit for the employees of IITM and Govt servants.

Last date for receipt of application is 29 December 1995.

Candidates already employed in Government/Semi-Government organizations/Public Sector Undertakings should apply through proper channel.

Special Drive Advt. No. Est. 4/95.

INDIAN COUNCIL OF AGRICULTURAL RESEARCH

The Indian Council of Agricultural Research, New Delhi announces the following existing and new Awards.

Rafi Ahmed Kidwai Awards, 1993–95

All research workers above 35 years of age engaged in the research in the field of agriculture, animal husbandry and allied sciences in India shall be eligible for the Award. Nine awards of Rs 100,000 each are to be given once in three years for the work done during the preceding three years of the period of the award.

Fakhruddin Ali Ahmed Awards, 1994–95

All research workers engaged in research in agriculture, animal husbandry, fisheries and allied sciences in tribal areas (tribal areas for this purpose are those notified by the Ministry of Home Affairs, Govt of India) will be eligible for applying for the Award. Two prizes of Rs 50,000 each are to be given once in two years for the work done during the preceding two years of the period of the award.

Vasant Rao Naik Award, 1995

All scientists or extension workers who have made outstanding contribution in the areas of water conservation and dryland farming in India shall be eligible for the award. One Award of Rs 100,000 is to be given annually, for the work done during the preceding five years.

Outstanding Teachers Awards, 1994–95

All staff members engaged in undergraduate or postgraduate teaching for at least 10 years in deemed universities, ICAR Institutes and State Agricultural Universities shall be eligible for the award. Eight awards of Rs 20,000 are to be given once in two years.

Outstanding Women Agricultural Scientist Award, 1995

All women scientists working at the Institutions under the ICAR system including State Agricultural Universities are eligible for the award. One award of Rs 25,000 to be given annually for the significant contributions made during her career as an agricultural scientist.

The application form for the above awards may be obtained from Dr R. C. Maheshwari, Asst Director General (CSC), ICAR, Krishi Bhawan, Dr Rajendra Prasad Road, New Delhi 110 001, on or before 15.2.1996 by sending a self-addressed one rupee stamped envelope. Six copies of the application may be forwarded by the Head of the Institution/Directors of Institutes/State Dept of Agriculture/Horticulture/Animal Science/Fisheries and Vice-Chancellors of State Agricultural Universities. The last date of submission of applications for these awards is 15 March 1996. The last date for candidates in the Andaman and Nicobar Islands, Lakshadweep States, Union Territory in the North Eastern Region, Ladakh Division of J&K State and Sikkim is 31 March 1996.

Edited and published by Prof. S. Ramaseshan and Prof. P. Balaram, Current Science Association, Bangalore 560 080.
Typeset by Creative Literati Pvt. Ltd., Bangalore (Tel. cum Fax: 5462698), Printed at Printek, Bangalore (3357763)

Current Science

Special Publications

Title	Guest Editor	Price India (Rs)	Price Abroad (US$)
1. *S. N. De and Cholera Enterotoxin*	P. Balaram	50	12
2. *The Ramachandran Map and Protein and Peptide Structures*	P. Balaram	50	12
3. *Waves and Symmetry*	G. Srinivasan	50	15
4. *A Celebration of Colour in Astronomy*	G. Srinivasan	175	20
5. *Clinical Research and Health Care in Developing Countries*	G. N. Menon	50	12
6. *Biotechnology in India* (copies out of stock)	P. Balaram	–	–
7. *Remote Sensing for National Development* (copies out of stock)	M. G. Chandrasekhar, B. L. Deekshatulu, V. Jayaraman, George Joseph, R. L. Karale, K. Kasturi-rangan, S. Krishnamurthy, K. Radhakrishnan	–	–
8. *Extinct Plants, Evolution and Earth's History*	B. S. Venkatachala, C. V. Subramanian	70	20
9. *Seismology in India – An Overview*	Harsh K. Gupta	100	20
10. *Alzheimer's Disease: An Emerging Issue for the Developing Countries*	George M. Martin	150	15
11. *Artificial Intelligence*	R. Narasimhan	100	12
12. *Quaternary Period in India*	K. S. Valdiya	150	20
13. *Pollination Biology in Tropics*	K. S. Bawa, K. N. Ganeshaiah, R. Uma Shaanker	100	12
14. *Special Issue on Inner Planets*	R. N. Singh	150	25
15. *Condensed Matter Science – Current Status and Plan for Action*	T. V. Ramakrishnan S. V. Subramanyam	150	20
16. *Neuroscience – A multidisciplinary approach*	P. N. Tandon Raghavendra Gadagkar	150	20
17. *Reproductive Biology*	T. C. Anand Kumar	150	20

Contact: The Circulation Department, Indian Academy of Sciences, P. B. No. 8005, Bangalore 560 080, India.

Tel.: 91-80-3342546; Grams: ACADEMY; Telex: 845-2178 ACAD IN; Fax: 91-80-3346094.

CURRENT SCIENCE

A fortnightly journal of research

Editors: S. Ramaseshan and P. Balaram

Editorial Board

S. Arunachalam, Central Electrochemical Research Institute, Karaikudi 623 006
Asis Datta, School of Life Sciences, Jawaharlal Nehru University, New Delhi 110 067
D. Balasubramanian, Centre for Cellular and Molecular Biology, Hyderabad 500 007
J. Chandrasekhar, Department of Organic Chemistry, Indian Institute of Science, Bangalore 560 012
R. Gadagkar, Centre for Ecological Sciences, Indian Institute of Science, Bangalore 560 012
K. N. Ganeshaiah, Department of Plant Genetics and Breeding, University of Agricultural Sciences, Bangalore 560 065
R. V. Hosur, Tata Institute of Fundamental Research, Bombay 400 005
N. V. Joshi, Centre for Ecological Sciences, Indian Institute of Science, Bangalore 560 012
Jyotirmoy Das, Indian Institute of Chemical Biology, Calcutta 700 032
C. C. Kartha, Division of Cellular and Molecular Cardiology, Sree Chitra Tirunal Institute for Medical Sciences and Technology, Thiruvananthapuram 695 011
S. S. Krishnamurthy, Inorganic and Physical Chemistry Department, Indian Institute of Science, Bangalore 560 012
S. C. Lakhotia, Department of Zoology, Banaras Hindu University, Varanasi 221 005
K. Muralidhar, Department of Zoology, Delhi University, Delhi 110 007
R. Narasimhan, CMC Limited, Bangalore 560 001
K. Neelakantan, ANURAG (Advanced Numerical Research and Analysis Group), Hyderabad 500 258
R. Nityananda, Raman Research Institute, Bangalore 560 080
T. Padmanabhan, Inter-University Centre for Astronomy and Astrophysics, Pune 411 007
B. R. Pai, Propulsion Division, National Aerospace Laboratories, Bangalore 560 017
T. J. Pandian, School of Biological Sciences, Madurai Kamaraj University, Madurai 625 021
Y. S. Rajan, Department of Science and Technology, New Delhi 110 016
M. Ramakrishnan, Geological Survey of India, Calcutta 700 016
A. V. Ramani, T. T. K. Pharma Ltd., Bangalore 560 025
S. Ramaswamy, Department of Physics, Indian Institute of Science, Bangalore 560 012
A. K. Raychaudhuri, Department of Physics, Indian Institute of Science, Bangalore 560 012
S. R. Shetye, Physical Oceanography Division, National Institute of Oceanography, Dona Paula, Goa 403 004
V. Siddhartha, 51 Bharati Nagar, New Delhi 110 003
V. Srinivas, School of Mathematics, Tata Institute of Fundamental Research, Bombay 400 005
G. Srinivasan, Raman Research Institute, Bangalore 560 080
R. Uma Shaanker, Department of Crop Physiology, University of Agricultural Sciences, Bangalore 560 065
K. S. Valdiya, Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore 560 064
G. Venkataraman, Department of Physics, Sri Sathya Sai Institute of Higher Learning, Prasanthi Nilayam 515 134
M. Vidyasagar, Centre for Artificial Intelligence and Robotics, Bangalore 560 001
K. VijayRaghavan, TIFR Centre, Indian Institute of Science Campus, Bangalore 560 012

Current Science, founded in 1932, is published by the Current Science Association in collaboration with the Indian Academy of Sciences. The journal is also intended as a medium for communication and discussion of important issues that concern science and scientific activity. All articles published in *Current Science*, especially editorials, opinions and commentaries, letters and book reviews, are deemed to reflect the individual views of the authors and not the official points of view, either of the Current Science Association or of the Indian Academy of Sciences.

Working Committee, Current Science Association

P. Balaram, Indian Institute of Science, Bangalore
R. Chidambaram, Department of Atomic Energy, Bombay
V. K. Gaur, CSIR Centre for Mathematical Modelling and Computer Simulation, Bangalore
K. Kasturirangan, ISRO Satellite Centre, Bangalore
P. K. Kaw, Institute for Plasma Research, Gandhinagar
N. Kumar, Raman Research Institute, Bangalore (*Vice-President*)
N. V. Madhusudana, Raman Research Institute, Bangalore (*Secretary*)
A. Mani, Field Research Unit, Indian Institute of Tropical Meteorology, Bangalore (*Past-President*)
R. A. Mashelkar, Council of Scientific and Industrial Research, New Delhi
R. Narasimha, Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore
G. Padmanaban, Indian Institute of Science, Bangalore (*Vice-President*)
R. Ramachandran, The Institute of Mathematical Sciences, Madras
P. Rama Rao, Defence Research and Development Organization, New Delhi
S. Ramaseshan, Raman Research Institute, Bangalore (*Treasurer*)
M. S. Valiathan, Manipal Academy of Higher Education, Manipal (*President*)

Editorial Office

Editorial staff

Chandrika Ramesh, G. Madhavan, T. D. Mahabaleswara, C. S. Ravi Kumar, M. S. Venugopal

Circulation and Accounts

Peter Jayaraj, Ranjini Mohan, B. Sethumani, Shanthi Bhaskar, B. K. Shivaramiah

Annual subscription 1995

	Institutional	Personal
India	Rs 350	Rs 100
All other countries (AIR MAIL)	US $175	US $50

Single copies other than special issues: Rs 35/US $15

Editorial Office: CURRENT SCIENCE
C. V. Raman Avenue
P.B. 8001
Bangalore 560 080, India
Telephone: 91-80-3342310
Telex 845-2178 ACAD IN
Fax: 91-80-3346094

Copyright © 1995, Indian Academy of Sciences

International Conference on

EUKARYOTIC EXPRESSION VECTOR SYSTEMS: BIOLOGY AND APPLICATIONS

February 4–8, 1996, New Delhi, India

Organizers: S. E. Hasnain, National Institute of Immunology, New Delhi; N. Khanna, International Centre for Genetic Engineering & Biotechnology, New Delhi; M. D. Summers, Texas A&M University, College Station (Chair).

The meeting will discuss state of the art in baculovirus, mammalian viruses, yeast, *Pichia* and other heterologous gene expression systems. Advancements in medical and veterinary therapeutics, biotech drug targets, diagnostics, vaccines and gene therapy, biopesticides, fermentation and downstream processing will be presented. The format will consist of invited lectures, short talks and poster sessions. The proceedings, after peer review, will appear in GENE.

Speakers: Mike Betenbaugh, Johns Hopkins University; David Bishop, IOV; Samir Brahmachari, Indian Institute of Science; Dipankar Chatterjee, Centre for Cellular & Molecular Biology; Asis Datta, JNU; Reiner Gentz, Human Genome Sciences, Inc.; K. P. Gopinathan, Indian Institute of Science; Seyed Hasnain, NII; Cornelis P. Hollenberg, Institut fur Mikrobiologei; Michael Karin, UC-San Diego; Navin Khanna, ICGEB; Jaspal Khillan, Thomas Jefferson University; Peter Liljestrom, Karolinska Institut; Susumu Maeda, UC-Davis; G. Padmanaban, Indian Institute of Science; Polly Roy, University of Oxford; Walter Schmidt, Research Institute for Molecular Pathology; Gale Smith, MicroGeneSys; Arun Srivastava, Indiana University; Koti Srikrishna, Marion Merrel Dow Inc.; Max Summers, TAMU; Waclaw Szybalski, UW-Madison; Inder Verma, Salk Institute; Just Vlak, Agricultural University; J. Wild, UW-Madison.

Fee of US$ 300.00 (Rs 3,000.00 for participants from India) will cover registration, lunch/dinner during the meeting. A few registration waivers available for deserving candidates. Send abstracts (15 × 20 cm) and brief CV before **December 22, 1995** to **Dr S. E. Hasnain, National Institute of Immunology, Aruna Asaf Ali Marg, New Delhi 110 067, India**; Tel: (91-11) 6863004 to 6863009, 6863011 to 6863013, 6862281; Fax; (91-11) 6862125, 6867626, 662608; Email: ehtesham@nii.ernet.in

Founded: 1932

ISSN: 0011-3891

CURRENT SCIENCE

Volume 69 Number 11

10 December 1995

891 **In this issue**

NEWS

893 **Hemp in the manufacture of paper**
S. Ramaseshan

CORRESPONDENCE

893 **Nurturing young scientists**
Umesh R. Desai

895 **Early research in India on primate tool using**
Srabani Sen and R. L. Brahmachary

895 **More on microbiology**
M. K. Chandrashekaran

896 **Karnatak music**
C. S. Unnikrishnan; J. R. Lakshmana Rao, *reply by* Raja Ramanna

SCIENTIFIC CORRESPONDENCE

898 **Too many hepatoprotective drugs and a little hepatoprotection**
A. Subramoniam; *reply by* R. T. Sane *et al.*

899 **Cryopreservation of epididymal spermatozoa**
N. Pandiyan, Radha Pandiyan, Jyotsna Malthai, P. S. R. Moorthy and B. Varalakshmi

900 **Cytokine antagonism by active vaccination**
Ruma Mukerjee and U. C. Chaturvedi

RESEARCH NEWS

903 **Driven up the tree: Why do wild dogs harass leopards?**
Anindya Sinha

GENERAL ARTICLES

904 **Some remarks on random numbers**
K. R. Parthasarathy

906 **Biodiversity conservation information network: A concept plan**
C. P. Geevan

914 **Earth's core**
S. K. Saxena

RESEARCH ARTICLE

918 **Effect of micellar and microemulsion media on the oxidation of iodide by peroxodisulphate reactions**
J. Santhanalakshmi and R. Kalaivani

RESEARCH COMMUNICATIONS

924 **Some FRW models with constant active gravitational mass**
Abdussattar and R. G. Vishwakarma

926 **Evaluation of intra-specific variability in *Avicennia marina* Forsk. using RAPD markers**
P. Balakrishna

929 **Carbon and oxygen isotope trends in late Precambrian–Cambrian carbonates from the Lesser Himalaya, India**
B. Kumar and V. C. Tewari

932 **Hypoglycaemic and antihyperglycaemic effects of *Aegle marmelos* leaves in rabbits**
V. V. Rao, S. K. Dwivedi, D. Swarup and S. R. Sharma

934 **Do dholes (*Cuon alpinus*) live in packs in response to competition with or predation by large cats?**
Arun B. Venkataraman

937 **Estimates of global methane production from rice paddies based on substrate requirement**
Sabyasachi Ghosh, M. C. Jain and Suresh K. Sinha

939 ***Bacillus thuringiensis*, a bicontrol agent for major tea pests**
N. Unnamalai and Vaithilingam Sekar

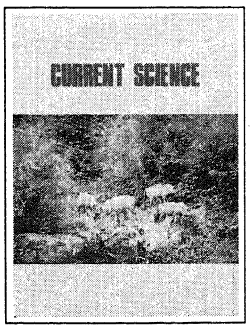
941 **Morphology of zircons in massive and incipient charnockites of southern Kerala: Their bearing on the origin**
G. R. Ravindra Kumar and Narayanaswamy

BOOK REVIEWS

945 **Annual Review of Immunology 1995,** *reviewed by* Satyajit Rath

946 **Annual Review of Medicine 1995,** *reviewed by* B. M. Hegde

946 **Tropical Diseases: Molecular Biology and Control Strategies,** *reviewed by* Manoj K. Das



COVER. Four dholes, comprising a sub-set of a study pack, living within the Mudumalai sanctuary, Tamil Nadu, India. See pages 903 and 934. [Photo courtesy: R. Sukumar]

Indexed in CURRENT CONTENTS/GEOBASE/CHEMICAL ABSTRACTS

In this issue

Random numbers

The notion of a 'random number' is of interest to scientists in a wide variety of disciplines, and is also the subject of philosophical speculations. This issue carries an article by K. R. Parthasarathy (**page 904**), an internationally renowned expert in probability theory, discussing some properties of random numbers.

It has been observed that population sizes of certain animals show peaks every 3 years, on the average. Parthasarathy's article starts with a quotation from C. R. Rao, which outlines a simple probabilistic explanation for this phenomenon. The article goes on to examine in more detail the following question; if $\xi_1, \xi_2, \ldots$ is a sequence of random numbers uniformly distributed in the unit interval [0, 1], what is the average length of time between successive 'k-peaks'? Here, ξ_n is a k-peak if it is greater than its k immediate predecessors or successors (that is, ξ_n is greater than any of $\xi_{n-k}, \xi_{n-k-1}, \ldots, \xi_{n-1}, \xi_{n+1}, \xi_{n+2}, \ldots, \xi_{n+k}$). The answer is shown to be $2k+1$, which for peaks (i.e. if $k=1$) reduces to 3, as pointed out by C. R. Rao. This is a consequence of an analysis carried out using the individual ergodic theorem of Birkhoff, and the Central Limit Theorem. The analysis also yields large sample tests to check if a given sequence of numbers is random.

V. Srinivas

Sample compositional model of the earth's inner core

The earth's core is the region extending from 2890 km to the centre at 6371 km. Several models exist to explain the composition of the earth's interior which is not available for direct observation but has to be inferred from a variety of disciplines. The outer part of the core is considered to be in liquid form since it does not transmit shear type of earthquake waves. The inner core is believed to be solid with a density of about 13.

It is commonly assumed that the core consists of iron and nickel, much the same composition as the meteorites. Various types of evidence point to iron being the major component of the core. Saxena (**page 914**) reviews some of the new techniques of heating solids with a laser at ultra high pressures in a diamond anvil. The new data lend support to a simple model of the core which is identical to that of iron and stony meteorites, the debris derived from differentiated broken-up planetary bodies. The density of such material seems to compare well with the seismic density. The reason why the outer core is in liquid form is explained to be the result of the phase transformation in iron and its melting behaviour.

B. P. Radhakrishna

VASVIK RESEARCH AWARDS

1995

OBJECTIVES

We have instituted seven annual VASVIK Industrial Research Awards, each of Rs. 25,000/-, a citation and a gold medal to be awarded to an individual or group of individuals, who, in the opinion of our Board of Advisers, have made outstanding contribution to the advancement of science and technology leading directly to national prosperity in India.

SCOPE

Consideration will be given to achievements in design or any production techniques or methods, particularly with regard to their impact on industrial and economic growth, import substitution, saving of foreign exchange, cost reduction, etc. Achievements within the past 10 years will be given priority, if they are of value to manufacturing industries, including the process of agricultural production.

FIELDS

The awards will be given to Indian citizens for research in each of the following fields:

a) Agricultural Sciences and Technology
b) Biological Sciences and Technology
c) Chemical Sciences and Technology
d) Electrical Sciences and Technology
e) Electronic Sciences and Technology
f) Material Sciences and Technology
g) Mechanical Sciences and Technology

SPECIAL AWARD FOR WOMEN SCIENTISTS

VASVIK has also instituted a special award solely for women scientists, known as SMT. CHANDABEN MOHANBHAI PATEL VASVIK AWARD FOR WOMEN SCIENTISTS, carrying a cash prize of Rs. 25,000/-, a citation and a gold medal, to recognise and encourage women's participation in science and technology leading to national prosperity.

NOMINATIONS

Applications/Nominations for the year 1995 may please be submitted in prescribed format before 30.4.1996

VIVIDHLAXI AUDYOGIK SAMSHODHAN VIKAS KENDRA
P.O. Box No. 9037, Patel Vanika
Western Express Highway, Goregaon (East) Bombay 400 063
Ph: 8401293, Fax: 91 (22) 8401298, 8401570
Telex: 1170035 – TUBE IN, Gram: "UDYOGVIKAS" Bombay 400 063

CURRENT SCIENCE

Volume 69 Number 11 10 December 1995

NEWS

Hemp in the manufacture of paper

Hemp – *Cannabis sativa* (often called Indian hemp) – has been cultivated in India for centuries, especially in regions where rivers flood their banks, because the plant holds the soil and prevents soil erosion. The plant produces strong pliable fibres (the word 'hemp' is used for both the plant and the fibres). The fibres are extremely strong and have been used since antiquity for weaving sail cloth, and for making large diameter ropes used in ships etc.

The male plant bears flowers in axillary racemes and dies soon after pollination has taken place. The female plant bears flowers in crowded spikes and dies after the seeds mature. Plants of both sexes are used for fibre. The hemp stem is hollow and has an inner fibrous bark. While strong coarse fibres are obtained from mature plants, soft ones obtained from hemp harvested at pollination are used for making cloth of fine texture. The fibres from the bark are used to make a great variety of textile products.

Hemp contains the psychoactive drug tetrahydrocannabinol, popularly known as marijuana. Narcotics like charas, ganja, etc. have been extracted in India from the hemp plant for centuries.

More recently, material scientists in India have considered hemp fibres as ideal for making fibre-reinforced plastic (FRP) composites because of the strength and length of the fibres. Further, hemp fibres contain a fair proportion of amorphous silica, making them behave almost like glass fibres except that hemp fibres are much tougher. Hemp fibres add strength and stiffness to plastic products and can be used to substitute glass (G) in GFRP products. Attempts to use treated hemp fibre to substitute partially asbestos in 'asbestos cement sheets' (used for roofing) have been fairly successful.

Now comes the news (*Environmental Health Perspective*, 1995, **103**, 893–894) that hemp is one of the best fibre sources for paper and therefore could be used in its manufacture. It is much better than jute, kenaf, flax or bagasse. In recent experiments, high-quality bond paper has been produced with 10% hemp together with recycled paper pulp. Lower grade but strong paper has also been produced for copiers. Hemp paper is considered environmentally very good as no chlorine bleach or acid treatment is necessary. Cultivating hemp, it is said, will save trees and minimize pesticide demands. Since cultivating the marijuana plant in the US is illegal, hemp pulp is being imported into the US from China and Hungary. Attempts are also being made to use genetic engineering to produce hemp not containing the drug tetrahydrocannabinol. There are some negative aspects too – Hemp plants require nitrogenous fertilizers; so hemp fibres may be much costlier compared to other materials like wood, bamboo, etc. used in paper manufacture. One does not know how the high silica content in this fibre will affect the paper-making machines.

S. Ramaseshan, Raman Research Institute, Bangalore

CORRESPONDENCE

Nurturing young scientists

It is said in common scientific parlance that best scientific contributions come from 'young' researchers. More often than not, Nobel laureates are rewarded for their work that was performed in their mid-thirties. Younger researchers are, hence, a breed that need to be nurtured with care all around the world and India, in particular. Given the right impetus, support, and direction this flock of assertive dedicated individuals can perform wonders. This is especially important in our changed times when India needs to compete and succeed globally on all frontiers of technology.

Much to the chagrin of young investigators, however, the research environment in our country is forsakenly taken to be rather dull and sleepy. A project proposal to government agencies, in general, takes anywhere from six months to more than a year for appraisal, to which six more months are to be added for money disbursal. The young scientist

proposal is expected to take much less; yet, invariably it consumes a lot of time. In these days of rapid technological change a loss of one year in research activity is like unquestionably accepting the forced directive of being condemned to technological backwardness, if not obsolesence. In much the same token, inefficient and inadequate financial support produces equivalent result. More important is the associated danger of a young scientist losing the self-confidence and the drive to compete internationally when saddled with these time and money problems at such an early stage of his career. If the nation needs technology to be delivered for social benefits, it is imperative that all researchers, and especially the younger ones, are supported duly. A nation is what its young people are and what they want to be.

The Department of Science and Technology operates a scheme for initial support of young scientists (below 35 years) to perform research. The basic idea is to provide some quick seed money required to establish a new laboratory/facility and allow the young investigator to take roots. At the present time this scheme can at best provide an amount of Rs 3 lakhs for a two-year period. At the earliest, the administrative machinery at DST can initiate the operation of a project under this scheme in a minimum of six months time; most projects usually take more. The *modus operandi* of initiating a project includes securing at least two referee reports. Recently, the Board of Research in Nuclear Sciences has come up with a similar scheme having a Rs 5 lakhs financial cap. It is anticipated that a BRNS young scientist project will take much shorter time to initiate.

The present manner of administering young scientist proposals is ineffective. The amount of time taken in processing the applications is much too large; money disbursal is evidently belated. I present some additions and modifications to the present set-up to help the young researchers' fraternity. Following are my point-wise recommendations with a brief discussion of appropriate rationale:

- *India should* **not** *have multiple Young Scientist schemes running in different government agencies.* If the goal of each scheme is the same, we need not have one scheme competing with the other. One manager is usually better than multiple managers—why duplicate effort?
- *The administration/management of young scientist proposals be taken over by previous Young Scientists.* A group of previous young scientists, who have experienced the process earlier, be nominated for handling the research proposals. The rationale is that these scientists will better appreciate the needs of their cadre and speeden-up the entire process. Governmental delays regarding appraisals, cataloging, communication, etc. will be reduced drastically. In addition, the applications be treated in the same manner as scientific papers to expedite the process. The process of project appraisal will essentially remain the same, however, communication with established scientists will be improved. Younger investigators at premier institutions such as Tata Institute of Fundamental Research, Indian Institute of Science, Indian Institute of Technology, etc., may be appointed for handling these proposals.
- *The time required to process a Young Scientist proposal be reduced to a maximum of three months.* This should include the review process, the decision on whether to fund, and communicating the final decision.
- *The recommendations of the reviewers be followed in budgetary allocation.* The young investigator be intimated of the budgetary allocation and asked for his confirmation or appeal. The approval of the young investigator on budgetary aspects should be deemed necessary for the final approval before the project is submitted to the finance manager for money remittance.
- *A finance manager(s) be appointed aside from the proposals management group who will handle the money disbursal and money utilization part of the proposal.* Ideally he should be a young scientist. If not, a cell in one of government bodies be designated to solely handle this function.
- *Money should be sent to the concerned investigator within two months of the approval of the proposal.* It is imperative that this function is given utmost importance. Time saved is money gained. If a cell or manager is identified with this function only, drastic reduction in time required may be achieved.
- *The maximum financial support be set to Rs 10 lakhs for a two-year period.* Out of this an investigator can claim up to 80% for the first year. This will allow greater latitude in research organization of the project and in management of facilities. A young investigator needs greater financial support earlier in his career than later; more so when he is setting up his laboratory.
- *In the first 3 years of initiating research activities, a young scientist may be allowed to work with two Young Scientist projects subject to the condition that he/she has no funding from any other source.* The rationale for this is the following: Very often in the initial stages one is not certain about success in a particular area and wants to explore more before committing himself/herself to a certain direction. In addition, better laboratory funding will allow greater flexibility in choosing advanced laboratory equipment that may be necessary for research.
- *The feasibility of processing applications before the actual employment should be maintained and encouraged.*
- *Under special circumstances a young scientist be allowed to present his proposal before the review committee without submission before-hand to expedite his application.* Some unforeseen events may delay the submission of applications. If appropriately justified, an application be considered before a meeting of the review committee for possible funding. This will help save more than a year for the young applicant.

The basic goal in all the above suggestions is to reduce the time required to initiate a project. If the process is considerably improved it may serve as an incentive for young researchers abroad to return to their motherland and do some good. A very encouraging thought is the reduction in brain-drain.

UMESH R. DESAI

Biotechnology Centre,
Indian Institute of Technology,
Bombay 400 076, India

Early research in India on primate tool using

Tool using by free ranging chimpanzees in East and West Africa is now well known to ethnologists. One of the earliest examples of the study of tool using by a captive chimp originates from the Teneriffe island research station[1]. This was followed by the famous work of Koehler[2] in 1915 and later summed up[3] in 1925.

We have recently come across a relatively obscure paper in which, as early as in 1902, tool using by a captive orang at Alipore Zoo, Calcutta was described by Sanyal[4], the first Indian Superintendent of that zoo. The orang at first tried, ineffectually, to reach the remnants of a proboscis monkey's breakfast in the adjoining cage. Then he collected straws from his sleeping quarters, fashioned a rough rope out of these. With the help of this rope he succeeded in pushing the tin of breakfast within reach.

1. Rothman, M. and Teuber, E., *Einzelausgabe aus der Anthropoidenstation auf Teneriffa*, 1915.
2. Koehler, W., *Aus der Anthropoidenstation auf Teneriffa*, 1915.
3. Koehler, W., *The Mentality of Apes*, New York, 1925.
4. Sanyal, R. B., *J. Asiatic Soc. Bengal (Calcutta)*, 1902, **LXXI(2)**, 92.

SRABANI SEN
R. L. BRAHMACHARY

History of Science Unit,
The Asiatic Society,
1, Park Street,
Calcutta 700 016, India

More on microbiology

R. Maheshwari's commentary 'Whither microbiology' (*Curr. Sci.*, 1995, **69**, 401–406) was an eye-opener even for a dyed-in-the-wool biologist like me. How many younger biologists would know without the benefit of this article that a study of microbiology will show that life is an autocatalytic process, that the principle of natural selection can be understood by raising enrichment cultures, that we can understand the unity of biochemistry, regulatory mechanisms, existence of permease, learn that bacteria invented photosynthesis and that plant chloroplasts are merely domesticated cyanobacteria, and that the study of *Neurospora crassa* led to the formulation of one-gene-one-enzyme hypothesis which revolutionized biology? Maheshwari has presented the case for the study of microbiology very persuasively and eloquently. To further strengthen Maheshwari's brief I quote the following from a recent article[1]: 'Recent analyses of phylogenetic frameworks have confirmed that the filamentous fungi share the closest evolutionary relationship to the Animalia of any non-animal group. As model systems for understanding biochemical genetics and gene regulation, the fungi have been particularly useful. Well-known examples include cracking the cell cycle using genetics and molecular biology, analysis of the mechanism of intracellular signaling via heterotrimeric G-proteins, cell type – specificity determination (all in the yeasts *Saccharomyces* and *Schizosaccharomyces*), and developmental patterning of growth habit as in the slime mould *Dictyostelium*. The best molecularly studied microbial clock system, and one of the best understood circadian systems in any context, is that of the ascomycete fungus *Neurospora crassa*, in which a circadian biological clock controls several aspects of growth and development. There is much reason to believe that the *Neurospora* clock system (input, mechanism and output) will be the first complete system to be resolved at the molecular level.'

Microbiology is nearly the whole of biology in a microcosm. Yet as C. Woese observes (cited by Maheshwari), 'University microbiology departments are withering. . . . Microbiologists are afraid to teach real microbiology anymore.' At the University of Wales, what was once a flourishing and world famous Department of Microbiology, headed by David Lloyd, was merged into rest of biology. I fervently hope that Maheshwari's stirring article will draw the attention of students who are on the threshold of choosing some kind of biology. In a lighter vein, in my ignorance I had all along believed that extremophiles, were scientists who went to extremes of maintaining that their kind of science alone mattered.

1. Loros, *J. Sem. Neurosci.*, 1995, **7**, 3–13.

M. K. CHANDRASHEKARAN

Dept of Animal Behaviour and Psychology,
School of Biological Sciences,
Madurai Kamaraj University,
Madurai 625 021, India

Karnatak music

I would like to point out that certain statements made in the article are not correct ('The structure of *Raga* music' by R. Ramanna, *Curr. Sci.*, 1995, **68**, 897–916). Ramanna states that a change of tonic (*Shruti*) in a piece of Carnatic music is strictly forbidden. This certainly is not true and there are musicians, especially the most skilled and imaginative, who use tonic shifts (*Shruti Bhedam*) routinely in their concerts. Examples can be quoted from numerous concerts of Balamurali Krishna and late Maharajapuram Santhanam and many others. This is done even in Carnatic instrumental music concerts. What is done is not a shift of the tuning of the *Tanpura*, but stressing the new *Sa*, which is actually any of the notes of the main *raga*, sufficiently to shift the mental tonic of the listener. It is also important to include the specific ornamentation of the derived *raga* as correct *gamakas* (which are in general different from the *gamakas* in the original *raga*) to get the desired effect. Simple examples are singing *ragas* like *Hindolam* and *Madhyamavati* from *Mohanam*.

A mere shift of the tonic without including the correct *gamakas* to change *ragas* in Carnatic music is not practical except for those *ragas* in which the *gamakas* are simple, essentially those that could be played on a harmonium or any keyboard instrument, and even there the effect is momentary, in those passages where even the *gamakas* have some similarity. This aspect cannot be seen in the written notations where any consistent mapping which preserves intervals more or less accurately looks viable. However, there are rare situations where the listener may hear a *raga* which is different from what the singer sings! This happens because the actual tuning of the *tanpura* is on notes *Sa*, *Pa* and *Sa*, whereas what the listener locks on to is *Pa* as his *Sa* and then the *Sa* as his *Ma* – a tonal shift from *Sa* to *Pa* in the drone itself. Then what he will perceive would be the inverted *raga* (in the sense of *raga* generated with the first inverted *shruti* chord). This happens in those situations where the listener is very familiar with the inverted *raga* and not the actual *raga* being sung. (I have heard it only in Hindustani music.)

Another important and controversial aspect which is touched upon (but elaborated only in the book) is the question of harmony in Indian music. R. Srinivasan in his review in the same issue has mentioned some of the practical difficulties in incorporating harmony while rendering Carnatic *ragas*. I merely want to stress the fact that harmony based on triads would be possible only in the obvious cases in which the *raga* has straightforward rendering without too many dynamical *gamakas* (roughly defined as modulations which blur the note in frequency all the way to the neighbouring notes, while stressing the main note itself. For example, many of the notes in *Todi*, combinations of *Ma*, *Pa* and *Dha* in *Kalyani*, *Panthuvarali* and derivatives, etc.) *Todi* would actually be very difficult in this respect with so many dynamical *gamakas* with the notes defined in a very impressionist way. But there are also many examples in which Indian musicians have tried out harmony with good success in *ragas* which do not depend strongly on dynamic *gamakas* (e.g. parts of *Mohanam* and many pentatonic *ragas*). There is a popular LP record by Ramani and his students on flute where many harmonized passages are played very skillfully, for different *ragas*. Harmony has been tried out many times in fusion music with a good Indian music component, and two-string harmony (may not really classify as harmony) has been tried by L. Shankar on some occasions (by recording twice over, playing his double violin) with pleasing effects.

There are statements in the beginning of the article which seem to suggest that only very recently Carnatic music has come out of courts into concert halls and that transformation to hi-fi era has not taken place in India. This obviously cannot be right. In fact, everything in hi-fi including Dolby technology, CDs, good acoustics, etc., are all incorporated already in many prominent recording studios as well as concert halls dealing with Indian music. So the justification that the article aims to educate technologists and scientists on aspects which might help in this technology revolution is misplaced, and after reading the article I also felt that it really does not contribute in this aspect even if there was a need.

The listing of the historical events of importance to Carnatic music (Appendix I) includes the name of Swati Tirunal, mentioned as a *Veena* expert and not as a composer! Is it a slip or taking sides in the well-known controversy?

While introducing Ramanna's article, the editor had commented that 'probably, many scientists may feel that it is not an appropriate subject for *Current Science*'. Many scientists would not really have any objection to publishing an article on music when it pertains to a scientific analysis or study of scientific aspects of music. In fact, some of the greatest of physicists have spent considerable time on musical acoustics and related theory. What was disappointing in the article was the way irrelevant information was included as transcripts of all the *Melakarta ragas* in western notation. This straightforward mapping, once the notation is established, should have been left as 'exercise' if anybody was interested in generating the full list.

C. S. UNNIKRISHNAN

Gravitation Experiments Group,
Tata Institute of Fundamental Research,
Homi Bhabha Road,
Bombay 400 005, India

Raja Ramanna's article fails to communicate and is thoroughly misleading.

In Part I of the article, an unsuccessful attempt is made to explain the notation used in Western music. As I could not follow it and as I have no knowledge of it myself, I refrain from commenting on what is presented in terms of that notation. As regards the rest of Part I, I say, without reservation, that it does not convey anything to the reader. In Part II of the article, there is a description of what the author considers to be the salient features of a typical Karnatak music concert. When one reads through it, one is amazed by its divergence from reality. For instance, mention is made of flutists playing accompaniment to vocal musicians in Karnatak music concerts. I regularly go to music concerts and, on the average, I have attended fifteen to

twenty concerts per year during the last fifty years and more. Not once have I found a flutist accompanying a vocalist.

Coming to percussion instruments played in Karnatak music concerts, mention is first made of *Tabla* which is seldom used in Karnatak music, and of morching which is rarely used. But the more commonly used *Ghata* and *Kanjira* do not find a mention anywhere.

While dealing with *alapana* it is stated that 'more often than not, the percussion instruments are silent during *alapana* as it does not always have a *tala* structure'. It means that *alapana* **has** sometimes a '*tala* structure', and on such occasions, percussion instruments **are** played. This is ludicrous, to say the least. An incorrect statement is also made about *tani avarthanam*: 'It may or may not be supported by other instrumentalists'. What sort of *tani avarthanam* is it if it 'is supported' by a non-percussion instrumentalist?

'After completing the "statement" of *pallavi* and *anupallavi*,' according to the author, 'the main musician takes on elaboration of the two themes in which new forces are brought in, such as *Kalpana swaras*'. *Kalpana swaras* being 'new forces' used to 'elaborate' *pallavi* is just absurd. Further, 'once *pallavi* and *anupallavi* is elaborated, there is a great degree of freedom for the performer for further elaboration swara-wise and tala-wise'. Statements of this sort which challenge our capacity for comprehension are abundant in the article.

'Two *ragas* with identical *Arohana* and *Avarohana* and derived from the same *Janya Mela* may yet differ from each other', says the author. Pray, what is *Janya Mela*? It is common knowledge that the 72 *Melakarthas* of Indian music are described as 'Sampurna' *ragas* because each of them have all the seven *swaras* in the order of their frequencies both in *Arohana* and *Avarohana*. They can also be called *Janaka ragas* because all the other *ragas* are derived from them. The derived *ragas* are *Janya ragas*. From this it is evident that *Janya Mela* is a contradiction in terms and so is meaningless. The context, however, suggests that *Melakartha* is meant and not *Janya Mela*.

As an instance of the above statement *Arabhi* and *Sama* are mentioned as two *ragas* with identical *Arohana* and *Avarohana* and derived from the same *Melakartha*. It is just not true. *Arabhi* is a derivative of *Dheerashankarabharana* and *Sama* of *Harikambhoji*. Their *Arohana* and *Avarohana* are also different.

J. R. LAKSHMANA RAO

'Maithri', XII Main,
Saraswatipura
Mysore 570 009, India

Raja Ramanna replies

I do not know the status of Lakshmana Rao as a musician and therefore if he is disappointed and thinks that there is nothing in my article, it cannot be quite correct because many of the things I have mentioned are based on well-known books on Karnatic music.

It is a pity, as Rao himself says that he does not know staff notations and therefore is unable to understand the comments made using these notations, and yet says that there is nothing in them. The article is written to make it comprehensible to the practitioners and scholars of western music as well as Indian counterparts. It has been necessary to do it through staff notation. Perhaps, the best thing for him to do, I feel, is to learn staff notation and then make his comments. A. M. Chinnasamy Mudaliar (1895), P. Sambamurthy and H. P. Krishna Rao of Mysore, etc. knew western notation and used it.

About the other points. The term Karnatic music concert has been used in a broad sense. In the music team of a Bharata Natyam performance, for instance, one can see the *Flute* functioning as a support to the vocalist. According to some old timers, before the advent of the *Violin*, *Veena* and or the *Flute* were the supporting instruments in any vocal recital. Many old timers in Mysore city are familiar with the flutist M. Venkatesa Iyengar (father of Doreswamy Iyengar) supporting Venkatagiriyappa in the *Veena* recitals. In recent times M. Balamuralikrishna has given many concerts with *Flute* support.

Regarding the use of *Tabla* in Karnatic music, any old timer in Mysore city will vouchsafe the fact that it was the *Tabla* (not *Mrudangam*) which was the accompanying talavadya in the music concerts before the *Mrudangam* appeared on the scene in the 1920s. It was similar in Bangalore city. About the absence of mention of *Ghatam*, *Khanjira* the relevant sentence in my article reads 'on his right side are placed percussion instruments like the *Tabla*, *Mrudangam*, *Morching* (Jews Harp), etc., covering other instruments instead of their detailed enumeration.'

About the *alapana*, Rao has laid emphasis on words 'more often than not' and draws his own sarcastic conclusion 'as if the *alapana* has some times a *tala* structure'. Here again, the use of the words *alapana* and *tala* are in a broader sense. Otherwise, how can one classify the *alapana* performed by nagaswaram where a person maintains a beat on the cymbals while the tavil also joins. I have had discussions with no less a person than Bangalore K. Venkataram, Director of the Percussive Arts Centre, Bangalore, about the *tala* structure in *alapana*. In fact, it was he who has suggested to me to include this particular aspect in my article. *Janya* instead of *Janaka Mela* is a typographical error.

Concerning other comments, answers can be found in Sambamurthy's volumes.

RAJA RAMANNA

National Institute of Advanced Studies,
Indian Institute of Science Campus,
Bangalore 560 012, India

Too many hepatoprotective drugs and a little hepatoprotection

This is in response to the Research Communication entitled 'Hepatoprotection by *Phyllanthus amarus* and *Phyllanthus debilis* in CCl_4-induced liver dysfunction' (R. T. Sane *et al.*, *Curr. Sci.*, 1995, **68**, 1243–1246) which describes the 'effectiveness' of *P. debilis* and *P. amarus* as hepatoprotective agents. The data are not presented properly and the conclusions drawn from them are partly misleading.

To evaluate hepatoprotection the data for CCl_4 + *P. amarus*-treated groups have to be compared with CCl_4-control group and not with normal control, and the statistical significance has to be determined between these groups. This has not been done. If there is complete hepatoprotection in CCl_4 + *P. debilis*- or *P. amarus*-treated groups, the values for serum enzymes and liver RNA shown in Tables 1 and 2 should not differ significantly between normal control group and CCl_4 + herbal drug-treated groups.

In Table 2, values for liver DNA and RNA are given. The levels of RNA decreased in the liver of CCl_4-treated animals compared to normal control. The RNA value for *P. amarus*-treated group (CCl_4 + *P. amarus*) is not statistically (significantly) different from that for CCl_4-control group. Furthermore, Table 2 has errors. All RNA values including the normal control value are marked with asterisks to indicate the level of significance. The SD for normal control group is 10.56 whereas that for group II is given as 0.7284. This is highly unlikely.

We have calculated the level of hepatoprotection in the herbal drug-treated groups based on the values given in Tables 1 and 2 for serum enzymes and liver RNA and presented in Table A. In the case of *P. amarus*, out of 3 serum marker enzymes studied, only GPT activity shows hepatoprotection (63%). The level of alkaline phosphatase activity does not show any protection. GOT activity shows a marginal effect (36%). Further, CCl_4-induced decrease in liver RNA content was not significantly influenced by *P. amarus* treatment. With these observations it should not be concluded that *P. amarus* is effective as a hepatoprotective agent. A good hepatoprotective agent must influence almost all the relevant biochemical parameters more or less uniformly. If not, a plant drug control group has to be included to find whether the plant drug influences any specific parameters in normal rats.

In *P. debilis*-treated animals some hepatoprotection is seen. Even in this case, GOT and alkaline phosphatase activities show only around 50% protection and liver RNA level shows only 28% protection. The normal recovery from damage without herbal drug treatment in three more days (six days after CCl_4 treatment) is far better than *P. debilis* treatment for three days. As compared to the values in CCl_4-control on day 3, on day 6 the recovery from damage in serum enzymes and RNA is 60–80%.

Although the authors compared all the biochemical parameters with those of the normal control rats, the normal control rat liver histology figure is not given. It is better to show that, at least, for the benefit of those readers who are not familiar with liver histology. Some variations in the degree of histopathological changes can be seen in different rats of the same group and also, minor variations can be observed in different parts of the same liver. When these points are considered, among other things, the presented minor improvements in histopathological changes in the CCl_4 + herbal drug-treated groups as compared to CCl_4-control group may not represent good hepatoprotection.

Thus the reported data show, at best, a marginal hepatoprotective effect in *P. amarus*-treated rats. Considering all the parameters studied, the hepatoprotective effect of *P. debilis* is also below 50%. A reasonably good hepatoprotective agent should exhibit, at least, 60–80% hepatoprotection as seen in the case of normal recovery, 6 days after CCl_4-treatment (Table A). Drugs exhibiting at least 60–80% protective effect, preferably 80–100% protection, may be considered as hepatoprotective agents for treatment in crude forms after establishing their safety.

Although *P. debilis* showed moderate/slight hepatoprotection it should be noted that this effect of *P. debilis* was observed at a dose which gave maximum effect (0.66 g drug powder/kg body wt). The dose response effect is very important, and the data for the same are not given. The effect may not be seen or deleterious effect may be present at higher doses. Caution is required to use those drugs which do not give a normal dose response curve. Arbitrary dose fixation in herbal-drug treatment can result in the loss of moderate/small hepatoprotective effect of *P. debilis*.

Many common hepatoprotective medicinal plants are shown to have more than 60% hepatoprotection in most of the relevant biochemical parameters. However, comparison with published literature is sometimes difficult owing to the difference in the level of liver damage induced by various hepato-toxic agents (CCl_4, paracetamol, galactosamine, etc.) and in treatment protocol. To determine effectiveness as a hepatoprotective agent it is always better to compare with a hepatoprotective drug known for its efficacy and/or formulations such as Liv52.

We have studied both *P. debilis* and *P. amarus* and could not find substantive hepatoprotective effects in these herbs. We have observed relevant serum

Table A. Hepatoprotective effect of *P. amarus* and *P. debilis* in rats (% hepatoprotection)

Parameters	Groups			
	CCl_4-control	CCl_4 + *P. amarus*	CCl_4 + *P. debilis*	CCl_4 (Normal recovery in 6 days)
GPT	0	63	98	60
GOT	0	36	45	65
Alk. P	0	8	57	80
RNA (liver)	0	12	28	67

Percentage of protection was calculated as follows, assuming that there was no protection (100% damage) in CCl_4-control group

$$\text{Percentage of protection} = 100 - \left(\frac{100}{\text{Difference in the values between } CCl_4 \text{ control and normal control}} \times \text{Difference in the values between } CCl_4 \text{ control and } CCl_4 + \text{drug group} \right).$$

enzymes, bilirubin, choleretic activity and liver histology to assess liver damage and protection from damage. In our studies on *P. debilis* we have found only 30–40% protection from paracetamol-induced elevation of serum enzymes and bilirubin at a dose of 1 g dry powder/kg body wt and we consider it as a small effect. In our screening we could detect *Phyllanthus* species which give 80–100% hepatoprotection regarding most of the biochemical parameters and liver histology.

There are numerous medicinal plants which have marginal or insufficient curative potentiality. Reporting them as effective or potent herbal remedies is misleading and dangerous to those who want to use them as remedies. Such reports will increase the numbers in the long list of hepatoprotective herbal drugs but not provide hepatoprotection to the patients.

A. SUBRAMONIAM

Ethnopharmacology Division,
Tropical Botanic Garden and
Research Institute,
Palode, Trivandrum 695 562, India

R. T. Sane et al.'s reply

The publication presented the preliminary findings of our studies on the two *Phyllanthus* sps. We have not claimed *Phyllanthus debilis* to be the best hepatoprotective agent. The publication was initiated because our surveys indicated that all species of *Phyllanthus* are being used as hepatotonics with no species considerations. The study has brought out clearly the species-dependent variation in the hepatoprotective action and highlights the need to be very selective in their use.

Statistical comparison of treated groups with the normal control group is justified because the extent of recovery after treatment will be better evaluated with normal liver rather than a damaged one, as suggested by A. Subramoniam (AS). Even if such a comparison is made, no significant changes will occur in our evaluation. We have not claimed complete hepatoprotection in our findings.

We regret inadvertent error of marking the normal values also with asterisks. The SD of group II is 0.7284 and has been reported correctly. Several papers have been published, which clearly show that *P. amarus* is not a good hepatoprotective agent. The suggestion about adding plant control group is good. We appreciate the analysis of percentage hepatoprotection for each of the experimental groups. It is evident that *P. debilis* is a better hepatoprotective agent than *P. amarus*. The recovery seen in normal recovery group is after a period of six days, whereas the recovery in plant-treated groups is after three days. Average recovery seen for each group from AS's analysis is given below:

CCl_4 control	CCl_4 + *P. amarus*	CCl_4 + *P. debilis*	Normal recovery
00%	30.0% after 3 days	57.0% after 3 days	68% after 6 days

It is very clear that the rate of recovery by natural regeneration is slower than the *P. debilis*-treated group. Besides, published papers indicate that complete recovery after CCl_4-induced damage is achieved only after two weeks. We do agree that there is better recovery of RNA levels in natural recovery group.

The photograph of normal liver histology was not provided since the structural characteristics needed no special mention. After scanning the entire liver area, the most representative area was reproduced and we do not consider the improvement seen in them as minor but, quite significant.

The dose of the plant slurry was not fixed arbitrarily. We have carried out a dose–response study, after which the specific dose of *P. debilis* was fixed. The details of these findings could not form part of a short research communication. We agree that dose fixation is very important to avoid doses of deleterious effects or no effect. In fact, we have also carried out toxicity studies on these plant slurries. In the light of the findings from these studies we found the dose of 1 g/kg to be on the higher side and mildly toxic. This could be the reason why experiments with *P. debilis* could not provide satisfactory results.

It is also to be noted that various hepatotoxicants cause hepatic injury in different ways. Therefore, evaluation of hepatoprotective or hepatocurative actions need to be carried out with careful analysis to avoid over-interpretation. We are eagerly awaiting the publication on screening experiments by AS, where a *Phyllanthus* sp. is said to be identified, with better hepatoprotective action.

We have already initiated further studies on other *Phyllanthus* sps. and plan to use different hepatotoxicants. We have also carried out electron microscopic studies and have evaluated several other biochemical parameters. We are confident of publishing our findings soon. We still believe *P. debilis* to be a good hepatotonic.

R. T. SANE
V. V. KUBER
MARY S. CHALISSERY
S. MENON*

Departments of Chemistry and
**Biological Sciences,*
Ramnarain Ruia College, Matunga,
Bombay 400 019, India

Cryopreservation of epididymal spermatozoa

In 1949, Polge *et al.*[1] in England made a serendipitous discovery that glycerol protected fowl spermatozoa from the otherwise lethal effects of freezing. In 1953, Bunge and Sherman[2] demonstrated for the first time that frozen and thawed human spermatozoa could result in pregnancy and the birth of normal babies. All subsequent published reports have confirmed the cryopreservability of ejaculated seminal spermatozoa. The first semen bank in India was started at Apollo Hospital in January 1988.

Obstructive azoospermia is an important cause of male infertility. Surgical correction of obstruction yields poor results except when single tubular anastamosis is undertaken[3]. In some cases, surgical

correction is not possible (as in vasal-aplasia). However, in most of these men, a good number of spermatozoa with variable motility could be obtained by epididymal aspiration. It is not always possible to use the spermatozoa obtained by epididymal aspiration for Assisted Reproductive Technology as this requires an additional expense and simultaneous preparation of the wife. Therefore it was decided to try and cryopreserve epididymal spermatozoa obtained from caput epididymis (in men undergoing scrotal exploration) so that this could be used at a later date if necessary.

The subjects for the study were three men undergoing scrotal exploration for obstructive azoospermia in 1991. Informed consent was obtained from all these men. Epididymal spermatozoa were aspirated from caput epididymis into 1 ml of Earle's balanced salt solution. The prefreeze spermiogram was done. This was then mixed with an equal volume of Matheson Carlborg–Gemzell medium (cryoprotectant[4]) and cryopreserved in a biological freezer at $-70°C$.

Good motile spermatozoa were obtained from all the three men from caput epididymis. The sperm concentration prefreeze were twelve million, ten million and thirty-five million per ml with 30%, 40% and 60% progressively motile spermatozoa. Vaso epididymal anastomosis was done on two men. One man had vasal aplasia. One of these men, two months later, showed a sperm concentration of eighty million per ml with 30% motility and his wife was pregnant a month later. She has since delivered a baby. The other man had no spermatozoa at the first follow up, a month later, and has not reported back since then. As these three men were not keen on Intra Uterine Artificial Insemination Husband or ART for their wives, the spermatozoa so obtained from the caput were frozen with their consent. After one month the sample was taken out and thawed and it was found that the spermatozoa in all the three samples had still maintained 40–50% of their prefreeze motility.

Surgical correction of obstructive azoospermia generally yields poor results (20%). Unless and until experience is gained with single tubular anastomosis, the results of vaso epididymal anastomosis are not likely to improve. However, many centres have good experience with GIFT/IVF by conventional means or by Micro Insemination and Sperm Transfer (MIST), Sub-Zonal Insemination (SUZI) or Intra Cytoplasmic Sperm Injection (ICSI). For all these techniques, the number of spermatozoa required is far less than for *in vitro* fertilization. Therefore, while attempting vaso epididymal anastomosis for obstructive azoospermia, if the couple are prepared, one could also undertake IUAIH or ART. If, however, the couple are not prepared, one could attempt to cryopreserve the epididymal spermatozoa and use the cryopreserved spermatozoa for ART later if the surgical correction were to be unsuccessful.

All mammalian cells are not cryopreservable. It is not known what offers the spermatozoa the cryopreservable capacity. The fact that epididymal spermatozoa could be cryopreserved indicates that human spermatozoa do not require the presence of seminal constituents for being cryopreserved.

Epididymal spermatozoa which are mature, motile and capable of fertilization are also cryopreservable. It is suggested that all attempts at surgical correction of obstructive azoospermia should be accompanied by an attempt at IUAIH/ART or at least cryopreservation of epididymal spermatozoa so obtained.

1. Polge, C., Smith, A. V. and Parkes, A. S., *Nature*, 1949, **164**, 666.
2. Bunge, R. G. and Sherman, J. K., *Nature*, 1953, **172**, 707.
3. Silber, S. J., *Fertil. Steril.*, 1978, **30**, 567–571.
4. Matheson, G. W., Carlborg, L. and Gemzell, C., *Am. J. Obstet. Gynaecol.*, 1969, **104**, 495–501.

N. PANDIYAN
RADHA PANDIYAN
JYOTSNA MALTHAI
P. S. R. MOORTHY
B. VARALAKSHMI

Department of Reproductive Medicine,
Apollo Hospitals,
Madras 600 006, India

Cytokine antagonism by active vaccination

The soluble mediators, known as cytokines, regulate immune response as well as effector phase of immune reactions. They are characterized by functional pleiotropy and redundancy and act as a 'network'. They are required for a number of normal reactions in the body but produce pathological changes when they are present at the wrong place and in wrong concentrations. Their role in a number of diseases has generated keen interest in the cytokine antagonists and their potential therapeutic use. The cytokine antagonists are nonpeptide molecules, soluble receptors, IL-1ra/ mutated cytokines and autoantibodies[1]. Besides these, the adverse effects of cytokines have been neutralized successfully by passive immunization of animals by anticytokine-antibodies[2–6]. On the other hand, the non-neutralizing type of antibodies contribute to a positive therapeutic response by better targeting of the cytokines to the appropriate cells[7,8]. We have successfully used a novel approach to actively immunize mice with a dengue virus-induced cytokine, the mouse Cytotoxic Factor (mCF), which acts as 'pathogenesis-related protein', to protect against the pathological effects[9]. We wish to draw attention to this potentially therapeutic approach of 'anti-disease vaccination'.

During dengue type 2 virus (DV) infection of mice, a unique cytokine mCF is produced by T lymphocytes. mCF is a 43 kD molecule on native PAGE and has an isoelectric point of pH 6.5. The sequence of 19 amino acids of the *N*-terminus of mCF (compared at Gen Data Base, Distributed Information Centre, Indian Institute of Science, Bangalore) differs from those of other cytokines and DV-specific proteins[10]. Northern blot tests done with the oligonucleotide probes derived from this sequence shows the presence of mRNA for mCF in the spleen cells of DV-infected mice. The cDNA library of mCF has been constructed in λgt 11 and expressed in *Escherichia coli*

(Chaturvedi *et al.*, unpublished data). mCF kills mainly H-2A negative macrophages and T helper cells etc. of a variety of animal species by inducing Ca^{2+} influx[11–14] and nitrite production (A. Misra, R. Mukerjee and U. C. Chaturvedi – communicated). mCF is capable of reproducing various pathological lesions that are seen in human cases of dengue haemorrhagic fever/shock syndrome (DHF/DSS) which include increased vascular permeability and damage to the blood–brain barrier[4,6] (Figure 1). Treatment of human blood leucocytes with mCF *in vitro* produces changes[15] that are described in cases of DHF[16].

Recently, presence of an mCF-like cytokine has been shown in the sera of the cases of DHF/DSS (ref. 17). This human CF (hCF) has been purified by chromatography and characterized (Chaturvedi *et al.* – unpublished data). Further, production of hCF by human peripheral blood CD4 positive T cells has been shown by *in vitro* stimulation by DV (ref. 18). T cell-enriched subpopulation could be stimulated directly to produce hCF but the yeild of hCF was enhanced in the presence of monocytes/macrophages (Mϕ). This shows that the effect of DV can be both direct or via monocytes. The Northern blot analysis using oligonucleotide probe prepared on the basis of amino-terminal sequence of mCF shows the presence of mRNA in T cells[18]. A number of characteristics of mCF are shared by hCF (Table 1), including molecular weight, reaction in Western blots and capacity to increase capillary permeability[19–22] (R. Mukerjee and U. C. Chaturvedi – communicated). This is an extremely important breakthrough in DHF research as mCF has been shown to be a 'pathogenesis-related protein' in mice.

Mice immunized with 5 μg of mCF mixed with Freund's incomplete adjuvant were challenged with 3 μg mCF at weekly intervals up to 48 weeks. Protection against the mCF-induced increase in capillary permeability and damage to the blood–brain barrier was studied in normal control and immunized mice. A complete protection is observed at the 4th week which persisted at higher levels up to the 16th week and then declined sharply. A breakthrough in the protection occurred with higher doses of mCF in a dose-dependent manner. mCF-specific antibodies are present in the sera obtained from the immunized mice as detected by ELISA and Western blot test and by neutralization of the cytotoxic activity of CF *in vitro*. The peak antibody titres correlated with peak protection. Intracerebral challenge with a lethal dose of DV prolongs mean survival time and delays onset of symptoms of sickness in immunized mice compared with normal mice, but the titre of the virus in the brain was similar[9]. Recently, the effect of various adjuvants, including Freund's incomplete adjuvant, alhydrogel, bacille Calmette-Guerin (BCG), normuramyl-L-*N*-methylalanyl-D-isoglutamine octalymide (compound No. 84/246, synthesized at the Central Drug Research Institute, Lucknow and kindly provided by K. B. Mathur) and tetanus toxoid (TT) on the immunogenicity of mCF and their optimum doses have been studied. The best effect is obtained with 84/246 followed by that with TT (ref. 20).

Paralysis and death of mice occur due to replication of DV in the neuron cells, while damage to the blood–brain barrier resulting in cerebral oedema and associated symptoms is caused by both virus and mCF. On the other hand, cerebral oedema is produced, but not the mortality, by mCF on intraperitoneal (i.p.) inoculation of mice with DV, as the virus does not enter the brain[6]. By mCF-immunization, neither the virus titre nor the DV-induced paralysis and death could be prevented, but the latter are significantly delayed by 3–9 days and the mCF-induced pathological lesions are prevented[9,20]. The immunization provides almost 90% protection against damage to the blood–brain barrier on challenging with DV i.p. Therefore, this approach may be very useful in cases where the virus does not enter the central nervous system (CNS). For example, in the cases of human dengue, signs and symptoms of cerebral oedema and rarely of encephalitis may be observed but the virus has never been detected in their CNS[6]. In spite of extensive work done to understand the pathophysiology and pathogenesis of DHF/DSS, the precise mechanism remains unclear[23], necessitating an indepth understanding of the pathophysiology of dengue syndromes to possibly open up new strategies to develop efficient vaccines.

Dengue illness is one of the most rapidly expanding diseases of the tropics, with over two billion people at risk of infection and millions of cases occurring every year[24]. The epidemics of dengue are widespread in this country and repeatedly occur year after year[25]. Dengue viruses are of four closely related types and yet each is sufficiently distinguishable from the others, such that sequential infection may lead to life-threatening haemorrhages and shock reactions. This has caused several problems in the development of a vaccine with the result that no vaccine against dengue is yet available for general use. The ultimate

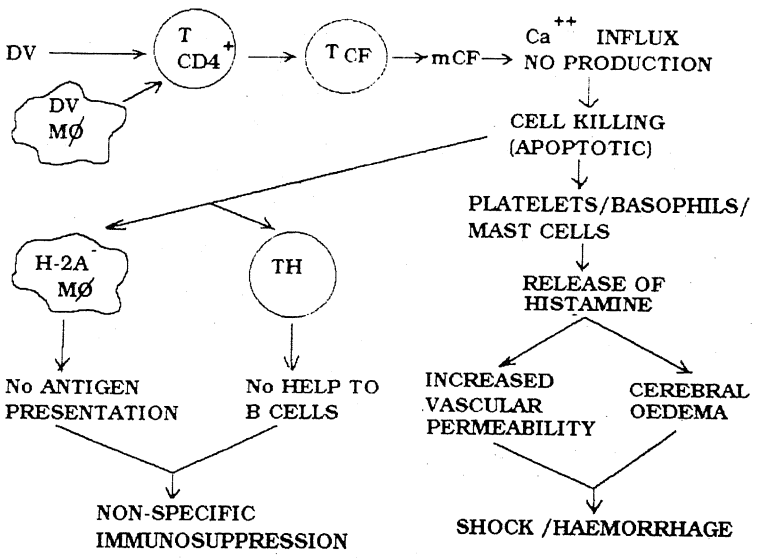


Figure 1. Production and mechanisms of action of mouse cytotoxic factor[19].

Table 1. Comparison of mouse (mCF) and human (hCF) cytotoxic factor[19]

Property	mCF	hCF
Inducer	DV[21]	DV[18]
Producer cell	L3T4 + T cells[21]	CD4 + T cells[18]
Target cell for assay	Mouse spleen cells[22]	Mouse spleen cells[18]
Mol wt on PAGE	43 kD (ref. 10)	43 kD (ref. 19)
Isoelectric point	6.5 (ref. 10)	4.5–5.2 (ref. 19)
N-Terminal sequence	19 amino acids known[4]	*N*-terminal blocked*
Effect on capillary permeability	Increased[4]	Increased†
Reaction with anti-mCF-antisera:		
– Cytotoxicity	Neutralized[4]	Neutralized*
– ELISA	Positive[17]	Positive[17]
– Western blot	Positive[10]	Positive[18]
– Dot blot	Positive*	Positive[18]
– mCF/hCF-induced increase of capillary permeability	Inhibited[4]	Inhibited†
Reaction with anti-hCF-antisera:		
– Cytotoxicity	Neutralized*	Neutralized*
– ELISA	Positive*	Positive*
– Western blot	Positive*	Positive*
– mCF/hCF-induced increase in capillary permeability	Inhibited†	Inhibited†
Northern blot‡	mRNA present in T cells*	mRNA present in T cells[18]

*Personal unpublished data.
†Mukerjee, R. and Chaturvedi, U. C. – communicated.
‡With the oligonucleotide probe prepared on the basis of amino-terminal sequence of mCF.

aim for dengue vaccine development should be the prevention and control of haemorrhages and shock which are life-threatening rather than the milder form of benign dengue fever. The data obtained so far indicate feasibility of developing an 'anti-disease vaccine'[9,20] for prophylaxis against DHF/DSS. On analysis of amino-terminal sequence of mCF[10] (at Gen Data Base, Bangalore) a projection of three antigenic sites was obtained on the molecule. Recombinant CF (rCF) has been obtained by expression of λgt 11 packed with cDNA in *E. coli* and some of its properties have been compared which show similarity between the native mCF and rCF (U. C. Chaturvedi – unpublished data). A question that needs to be addressed is whether such a vaccine could ever find practical use because it is based on immunization against a self-protein. Presence of autoantibodies to a number of cytokines have been described and their role were discussed[1]. Anticytokine autoantibodies are known to block the activity of the cytokines and are found in the serum for interleukin-1a (IL-1a), IL-2, IL-4, IL-6, IL-8, IL-10, TNF-α and IFN-α. Small complexes between such antibodies and cytokines do not activate complement if IgG4 antibodies are involved, and do not precipitate *in vivo*, but are active as inflammatory complexes, therefore, their therapeutic use should be considered carefully[1]. Bendtzen *et al.*[7] have proposed that a major role of anticytokines autoantibodies is to facilitate, rather than neutralize, functions of cytokines in the body by acting as specific physiological carriers and regulators of cytokines. For effective immunization it is essential that such epitopes on cytokines are used which generate neutralizing antibody only. This approach has the advantages and disadvantages of an active vaccination, but vaccine strategies directed at the primary cause (the cytokine) rather than infective agents need serious consideration, more so when effective vaccines are not available for several agents.

1. Debets, R. and Savelkoul, H. F. J., *Immunology Today*, 1994, **15**, 455–458.
2. Beutler, B., Milsark, I. W. and Cerami, A. C., *Science*, 1985, **229**, 869–871.
3. Digiovine, E. S. and Duff, G. W., *Immunology Today*, 1990, **11**, 13–20.
4. Khanna, M., Chaturvedi, U. C., Sharma, M. C., Pandey, V. C. and Mathur, A., *Immunology*, 1990, **69**, 449–453.
5. Dhawan, R., Khanna, M., Chaturvedi, U. C. and Mathur, A., *Int. J. Exp. Pathol.*, 1990, **71**, 83–88.
6. Chaturvedi, U. C., Dhawan, R., Khanna, M. and Mathur, A., *J. Gen. Virol.*, 1991, **72**, 859–866.
7. Bendtzen, K., Svenson, M., Jonsson, V. and Hippe, E., *Immunology Today*, 1990, **11**, 167–169.
8. Dhawan, R., Chaturvedi, U. C., Chaturvedi, P., *et al.*, *Int. J. Exp. Pathol.*, 1993, **73**, 455–461.
9. Chaturvedi, U. C., Mukerjee, R. and Dhawan, R., *Clin. Exp. Immunol.*, 1994, **96**, 202–207.
10. Khanna, M. and Chaturvedi, U. C., *Int. J. Exp. Pathol.*, 1992, **73**, 43–49.
11. Chaturvedi, U. C., *Immunology Today*, 1986, **7**, 159.
12. Chaturvedi, U. C., Dalakoti, H. and Mathur, A., *Immunology*, 1980, **41**, 387–392.
13. Chaturvedi, U. C., Nagar, R., Gulati, L. and Mathur, A., *Immunology*, 1987, **61**, 297–303.
14. Khanna, M., Chaturvedi, U. C., Dhawan, R., Tekwani, B. L. and Pandey, V. C., *Immunology*, 1991, **72**, 73–78.
15. Chaturvedi, U. C., Gulati, L. and Mathur, A., *Indian J. Med. Res.*, 1982, **75**, 469–473.
16. Wells, R. A., Scott, R. M. C. N., Pavanand, K., *et al.*, *Infect. Immun.*, 1980, **80**, 428–433.
17. Mukerjee, R. and Chaturvedi, U. C., *Indian J. Med. Res.*, 1995, **102**, in press.
18. Mukerjee, R., Chaturvedi, U. C. and Dhawan, R., *Clin. Exp. Immunol.*, 1995, **102**, 262–267.
19. Chaturvedi, U. C., Dhawan, R. and Mukerjee, R., in *Dengue and Dengue Hemorrhagic Fever* (eds Gubler, D. J. and Kuno, G.), CAB International, Wallingford, Oxon, 1995, in press.
20. Mukerjee, R. and Chaturvedi, U. C., *Clin. Exp. Immunol.*, 1995, **102**, 496–500.
21. Chaturvedi, U. C., Bhargava, A. and Mathur, A., *Immunology*, 1980, **40**, 665–671.
22. Chaturvedi, U. C., Mathur, K. R., Gulati, L. and Mathur, A., *Immunol. Lett.*, 1981, **3**, 13–16.
23. Halstead, S. B., in *Monograph on Dengue/Dengue Haemorrhagic Fever*, World Health Organization, SEARO, New Delhi, 1993, no. 22, pp. 80–103.
24. Koko, U., in *Monograph on Dengue/Dengue Haemorrhagic Fever*, World Health Organization, SEARO, New Delhi, 1993, no. 22, p. VII.
25. Banerjee, K., *ICMR Bull.*, 1994, **24**, 59–62.

ACKNOWLEDGEMENTS. The studies described here have been financed by the ICMR, New Delhi; CSIR, New Delhi and the UP State Government.

RUMA MUKERJEE
U. C. CHATURVEDI

Postgraduate Department of Microbiology, K. G. Medical College, Lucknow 226 003, India

Driven up the tree: Why do wild dogs harass leopards?

Anindya Sinha

A pack of Asiatic wild dogs or dholes suddenly appear on a grassy knoll in the darkening forests of the Mudumalai Sanctuary in southern India. They whine, sniff, rub against each other and then, as fleetingly, vanish into the dusk.

Why do carnivores like the dhole live out ardently social lives—hunting, feeding and resting together? Various selective advantages have been believed to accrue to communally living carnivores like wild dogs, lions and cheetahs. These include an increased likelihood of defending a territory with its associated resources of food and mates, greater success through cooperative hunting, and a better ability to protect kills from other predators. In an article in this issue, however, Arun Venkataraman provides tantalizing evidence for yet another kind of advantage that wild dogs, living in large cooperative groups, may enjoy[1]. He describes four instances where two of his study packs, which he has been studying continuously for five years now, actively mobbed leopards by grouping together and surrounding the trees in which the latter had taken refuge. In contrast, however, on a single observed occasion, the dogs bunched up and retreated from a tiger. This, together with the observation that leopard faeces contain more dhole hair than do those of tigers, leads Venkataraman to speculate that the wild dogs may be harassing leopards to deter them from preying on dogs isolated from the pack during hunting expeditions or on unprotected pups left behind in dens.

A competing hypothesis to explain the agonistic interactions of dholes preferentially towards leopards would suggest that the leopard's prey significantly overlaps with that of the dogs, while the prey of the tiger does not; competition for similar food may thus lead dholes to drive away leopards from their home range. In order to strengthen his own viewpoint, Venkataraman argues against this hypothesis rather convincingly. He shows that not only is there considerable overlap between the diets of the dhole–leopard, dhole–tiger and leopard–tiger pairs of carnivores, but the overlap is rather similar for the three pairs. Further support against the hypothesis comes from yet another interesting angle. Although leopards and dholes compete for similar-sized prey, there appears to be a niche-partitioning between the two predators based on the age- and sex-differences of the prey species. Thus, the dogs appeared to prey more on fawns or aged chital individuals; there was no such age-based selection by leopards. Moreover, dholes had a preference for male chital which the leopard did not. In conclusion, one important reason why dholes do live in large packs may therefore be to deter predation by leopards – surely a refreshing change from our familiar everyday dog-chase-cat situation!

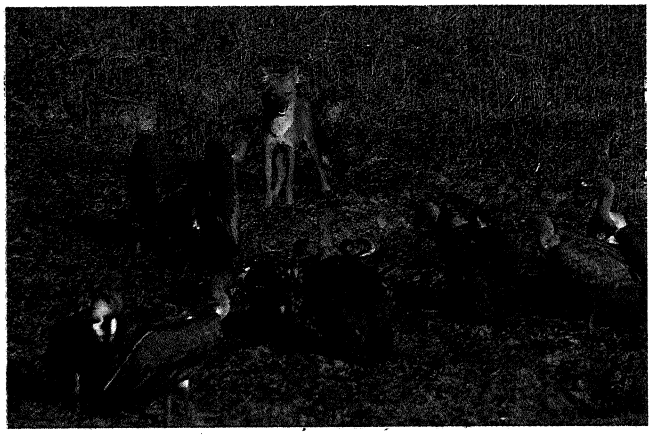


Figure 1. Dhole tensely approaching chital, killed by its pack earlier. Competition between vultures and the dholes is evident. (Photo: R. Sukumar).

The excitement of academic progress in any field of study can be easily judged from the gamut of refreshing ideas that challenge earlier theories and open up new avenues of research. Venkataraman's study is a valuable addition to studies on the behavioural ecology of social carnivores. Can other ideas be envisaged as one observes a pack of wild dogs mobbing a leopard that has been driven up a tree? Do all individuals participate in these attacks with equal enthusiasm? Are there dogs who cheat the rest of the pack by slinking away and leave the risky task of taking on the leopard's might to others? If indeed some individuals do so, are the other members of the group aware of them? Is there retribution for such cheating? Or do individuals move away boldly and are not punished since they serve the pack in other ways – by being good hunters or able defenders of the pups?

We now move away from a scenario where genetically related animals in a pack gain indirect fitness by helping one another. Is it possible that the dogs may potentially further their own individual fitness by displaying before each other their prowess in challenging a formidable adversary (V. Nanjundiah, pers. commun.)? Could this be a testing ground where individuals evaluate each other's strengths and capabilities as they plan future coalitions, challenges to the existing dominance hierarchy or even the choosing of mates? There is now increasing evidence that in many social systems such as those of carnivores and primates, individuals may be more aware of each other's strengths and weaknesses than was previously believed. In a recent study on African lions, for example, it has been shown that in coalitions of two individuals setting out to defend their territory from invading foreign animals, the more active individuals evaluate and appear to be aware of certain companions who persistently lag behind during such risky operations[2].

But given its scope and the fact that such detailed studies on the behavioural ecology of Indian wildlife are rare, there is no doubt that Venkataraman's study should provide a number of exciting answers in the years to come.

1. Venkataraman, A. B., *Curr. Sci.*, 1995, **69**, 934–936, (this issue).
2. Heinsohn, R. and Packer, C., *Science*, 1995, **269**, 1260–1262.

Anindya Sinha is in the National Centre for Biological Sciences, TIFR Centre, P.O. Box 1234, IISc Campus, Bangalore 560 012, India.

Some remarks on random numbers

K. R. Parthasarathy

It is shown that the number of k-peaks in a sequence of n random numbers and the time at which a k-peak occurs for the nth time in an infinite sequence of random numbers are asymptotically normally distributed. Exact expressions for their means and variances are evaluated.

IN his twenty-first convocation address 'Uncertainty, randomness and creativity' delivered at the Indian Statistical Institute, Calcutta on 5th March 1987, C. R. Rao[1] brings to our attention, in his characteristic style, an attractive feature of random numbers:

'It is an interesting property of random numbers that, like the Hindu concept of God, it is patternless and yet has all the patterns in it. That is, if we go on generating strictly random numbers we will encounter any given pattern sometime or other. Thus, if we go on tossing a coin, we should not be surprised if 1000 heads appear in successive tosses at some stage. So we have the proverbial monkey which, if allowed to type continuously, can produce the entire works of Shakespeare in a finite though a long period of time. . . .'

Continuing in this vein he deals with the following example:

'It is found that population sizes of a large variety of animals exhibit roughly a three-year cycle, i.e., the average time that elapses between two successive peak years of population size is about three years. (A peak year is defined as a year in which there are more animals than in the immediately preceding and immediately succeeding years.) The ubiquity of such a phenomenon led some to believe that perhaps a new law of nature has been uncovered. The belief was dealt a mortal blow when it was noted that if one plots random numbers at equidistant points, the average distance between peaks approaches 3 as the series of numbers gets large. In fact, such a property is easily demonstrable by using the fact that the probability of the middle number being larger than the others in a set of 3 random numbers is $\frac{1}{3}$. This gives an average distance of 3 years between the peaks.'

Here we elaborate on these comments by presenting a more detailed but elementary analysis of the frequency of occurrence of peaks (of arbitrary order) in a sequence $\{\xi_n\}$, $n = 1, 2, \ldots$ of independent and identically distributed random variables with a continuous distribution function $F(x)$. We may and do, without loss of generality, assume that ξ_n is uniformly distributed in the interval $[0, 1]$. This can always be achieved by considering the transformed sequence $\{F(\xi_n)\}$. We would consider this as a good pedagogical illustration of the individual ergodic theorem of Birkhoff and the central limit theorem for an m-dependent stationary sequence of random variables[2].

Say that a time point n is a k-peak for the sequence $\xi_1, \xi_2, \ldots$ if the event

$$E_n = \{\xi_n > \max(\xi_{n-k}, \xi_{n-k+1}, \ldots, \xi_{n-1}, \xi_{n+1}, \xi_{n+2}, \ldots, \xi_{n+k})\} \quad (1)$$

occurs. Clearly the probability of E_n is given by

$$P(E_n) = \int_0^1 x^{2k}\, dx = \frac{1}{2k+1}. \quad (2)$$

For $k = 1$, this is $\frac{1}{3}$ as mentioned earlier.

Denote by S the shift transformation defined by $S\xi_n = \xi_{n+1}$, $n = 1, 2, \ldots$. The distribution of the i.i.d sequence $\xi_1, \xi_2, \ldots$ is invariant and ergodic under S. For any event E denoted by 1_E its indicator which assumes the value 1 if E occurs and 0 otherwise. Then

$$\pi_{n,k} = 1_{E_{k+1}} + 1_{E_{k+2}} + \ldots + 1_{E_{n+k}} = 1_{E_{k+1}} + 1_{E_{k+1}} oS + 1_{E_{k+1}} oS^2 + \ldots + 1_{E_{k+1}} oS^{n-1} \quad (3)$$

is the number of k-peaks in the finite length sequence $\xi_1, \xi_2, \ldots, \xi_{n+2k}$. Note that the first likely k-peak in this sequence cannot be before the time $k+1$ and similarly the last likely k-peak cannot be after the time $n+k$. It now follows from the individual ergodic theorem of Birkhoff that

$$\lim_{n\to\infty} \frac{\pi_{n,k}}{n} = \frac{1}{2k+1} \text{ a.s.} \quad (4)$$

In other words the proportion of the number of k-peaks

K. R. Parthasarathy is at the Indian Statistical Institute, 7, S. J. S. Sansanwal Marg, New Delhi 110 016, India.

in a sequence of $n+2k$ random numbers converges almost surely to $1/(2k+1)$ as $n \to \infty$.

We shall now evaluate the variance of $\pi_{n,k}$. First observe that $1_{E_{k+r}}$ and $1_{E_{k+s}}$ are independent random variables whenever $|s-r| > 2k$. The covariance between $1_{E_{k+r}}$ and $1_{E_{k+s}}$ depends only on $s-r$ (and k which is fixed). The variance of $1_{E_{k+1}}$ is

$$\frac{1}{2k+1}\left(1-\frac{1}{2k+1}\right) = 2k(2k+1)^{-2}.$$

Thus

$$\operatorname{var}(\pi_{n,k}) = \frac{2nk}{(2k+1)^2} + 2\sum_{1 \le r \le 2k}(n-r)\operatorname{cov}(1_{E_{k+1}}, 1_{E_{k+r+1}}), \quad (5)$$

where 'var' denotes variance, and 'cov' the covariance. If $1 \le r \le k$ it is clear that $E_{k+1} \cap E_{k+r+1} = \phi$.

Thus

$$\operatorname{cov}(1_{E_{k+1}}, 1_{E_{k+r+1}}) = -\frac{1}{(2k+1)^2} \quad \text{if} \quad 1 \le r \le k. \quad (6)$$

For $k < r \le 2k$ we have

$$\begin{aligned}
&\mathbb{E}\, 1_{E_{k+1}} 1_{E_{k+r+1}} \\
&= P(\xi_{k+1} > \max(\xi_1, \xi_2, \ldots, \xi_k, \xi_{k+2}, \xi_{k+3}, \ldots, \xi_{2k+1}), \\
&\qquad \xi_{k+r+1} > \max(\xi_{r+1}, \xi_{r+2}, \ldots, \xi_{r+k}, \xi_{r+k+2}, \ldots, \xi_{r+2k+1})), \\
&= \mathbb{E}\,(\xi_{k+1}\xi_{r+k+1})^{r-1}\,(\xi_{k+1} \wedge \xi_{k+r+1})^{2k+1-r} \\
&= 2\int_{0<x<y<1} x^{2k} y^{r-1}\, dx\, dy \\
&= \frac{2}{(2k+1)(2r+1)}, \quad (7)
\end{aligned}$$

where $\wedge$ denotes minimum. Substituting (6) and (7) in (5) we conclude that

$$\operatorname{var}(\pi_{n,k}) = \frac{-2nk}{(2k+1)^2} + \frac{2k}{2k+1} + \frac{4}{2k+1}\sum_{r=k+1}^{2k}\frac{n-r}{2k+r+1}. \quad (8)$$

This implies

$$\lim_{n\to\infty}\frac{1}{n}\operatorname{var}(\pi_{n,k}) = \frac{2}{(2k+1)^2}\sum_{r=k+1}^{2k}\frac{2k+1-r}{2k+1+r} = \sigma_k^2, \text{ say.} \quad (9)$$

The following table gives the values of σ_k, $1 \le k \le 5$

k	1	2	3	4	5
σ_k	0.21082	0.16997	0.14517	0.12858	0.11655

By the central limit theorem for an m-dependent stationary sequence of random variables (see for example Theorem 19.2.1, ref. 2) we now conclude that

$$\left[\left(\pi_{n,k} - \frac{n}{2k+1}\right)\Big/(\sqrt{n}\,\sigma_k)\right] \overset{\text{in law}}{\Rightarrow} N(0, 1), \quad (10)$$

as $n \to \infty$, where $\Rightarrow$ denotes weak convergence and $N(0, 1)$ denotes the standard normal distribution with density function $(2\pi)^{-1/2}\exp -\frac{1}{2}x^2$. It may be noted that for large k, σ_k is approximately

$$\left\{\frac{2}{2k+1}\left(2\log\tfrac{4}{3} - \tfrac{1}{3}\right)\right\}^{1/2} = 0.38824(2k+1)^{-1/2}.$$

Denote by $\tau_{m,k}$ the time at which the mth k-peak occurs. Then

$$\begin{aligned}
P(\tau_{m,k} \ge n) &= P(1_{E_{k+1}} + 1_{E_{k+2}} + \ldots + 1_{E_{k+n}} \le m) \\
&= P(\pi_{n,k} \le m) \\
&= P\left[\left(\pi_{n,k} - \frac{n}{2k+1}\right)\Big/\sqrt{n}\,\sigma_k \right. \\
&\le \left.\left(m - \frac{n}{2k+1}\right)\Big/\sqrt{n}\,\sigma_k\right]. \quad (11)
\end{aligned}$$

Putting

$$\left(m - \frac{n}{2k+1}\right)\Big/\sqrt{n}\,\sigma_k = \alpha$$

and solving for n we get

$$n = (2k+1)^2/4\ \{2\alpha^2\sigma_k^2 + [4m/(2k+1)] - 2\alpha\sigma_k(\alpha^2\sigma_k^2 + [4m/(2k+1)])^{1/2}\}.$$

From (10) and (11) we now conclude that

$$\lim_{m\to\infty} P\left(\frac{\tau_{m,k} - m(2k+1)}{\sqrt{m}\,\sigma_k(2k+1)^{3/2}} > -\alpha\right) = \frac{1}{\sqrt{2\pi}}\int_{-\alpha}^{\infty} e^{-x^2/2}\, dx.$$

In other words

$$\sqrt{m}\,\frac{(\tau_{m,k}/m) - \overline{2k+1}}{\sigma_k(2k+1)^{3/2}} \overset{\text{in law}}{\Rightarrow} N(0, 1) \quad (12)$$

as $m \to \infty$. Roughly speaking the average length of the

time interval between the two successive k-peaks is $2k+1$. When $k=1$ this is 3 as pointed out in the beginning.

It is not difficult to see that the joint law of

$$\left[\sqrt{m} \left(\frac{\tau_{m,j}}{m} - \overline{2j+1} \right), \; j = 1, 2, \ldots, k \right]$$

converges to a k-dimensional normal distribution with a nondiagonal covariance matrix as $m \to \infty$. The asymptotic results (10) and (12) provide a series of large sample tests for testing the hypothesis that a sequence of numbers is random.

1. Rao, C. R., Uncertainty, Randomness and Creativity, 21st convocation address, Indian Statistical Institute, Calcutta, 1987.
2. Ibragimov, I. A. and Linnik, Yu. V., *Independent and Stationary Sequences of Random Variables*, Wolters-Noordhoff Publishing, Groningen, 1971, pp. 369–370.

Biodiversity conservation information network: A concept plan

C. P. Geevan

A network that ensures the availability of reliable, up-to-date environmental information is necessary to realize the objectives set out in Convention on Biodiversity (now a treaty) that followed from the Earth Summit at Rio, June 1992. The task of building a nature conservation information network should, therefore, be considered an important part of the biodiversity conservation agenda. This paper presents an outline of an hypothetical information network, designated as Conservation Information Network (CiNet), to meet requirements of bioresources conservation, mapping, inventorying and monitoring on a large scale. The dataflow framework presented takes into account the existing data networks in India.

Background

The Convention on Biological Diversity (CBD) that emerged from the United Nations Conference on Environment and Development (UNCED) or the Earth Summit at Rio de Janeiro, in June 1992 is now a treaty. The CBD covers almost every aspect of biodiversity conservation. Article 17 of CBD concerns exchange of information. However, it does not lay down any operational framework for achieving information exchange. Nevertheless, an information network that ensures the availability of reliable, up-to-date environmental information is necessary to realize the objectives set out in CBD. The existing information systems are considered to be inadequate to meet these challenges[1].

The task of building a nature conservation information network should, therefore, be considered an important part of the conservation agenda. This paper presents an outline of an information network to meet requirements of bioresources mapping, inventorying and monitoring programme. The hypothetical network is designated as Conservation Information Network (CiNet). The dataflow framework is presented taking into account the ready availability of well-developed data networks in India and the CiNet is conceptualized as an overlay network riding over the existing networks.

C. P. Geevan is in the Salim Ali Centre for Ornithology and Natural History, Kalampalayam P.O., Coimbatore 641 010, India.

Special interest groups on global networks

Computer networking has gone beyond setting up data links to creating information highways over which organizations and individuals are connected across the globe. The nature conservation efforts need to take advantage of these developments in information technology and create a niche for itself in the cyberspace. There are several important initiatives in this direction such as the INFOTERRA of the United Nations Environment Programme (UNEP) consisting of 170 national nodal points coordinated from the UNEP headquarters at Nairobi, the Environmental Resources Information Network (ERIN) in Australia with a biodiversity information system designed to meet the changing user needs, the Bio-diversity Information Network (BIN21) dedicated to the CBD with its secretariat at the Tropical Database

in Brazil, the Long-term Ecological Research Network (LTER) based in the University of Washington and the projects in Capacity Building for Biodiversity Information Management begun by World Conservation Monitoring Centre (WCMC) based in the United Kingdom and the BioNET-International proposed by CAB International to pool the global resources in biosystematics. With the assistance from the United Nations Development Programme (UNDP) under CAPACITY 21, a large number of Sustainable Development Networks are also expected to be set up.

The capacity building project of WCMC is supported by the European Union and allows it to carry out the strategic development work as well as testing of networking and other cooperative approaches for biodiversity information management. BioNET-International is a Global Technical Cooperation Network of institutions and people concerned with biosystematics of invertebrates and microorganisms. The goal of this network, as stated, is the mobilization and enhancement of the world's biosystematic resources for the benefit of developing countries. BioNET will build and sustain biosystematic self-reliance in developing sub-regions and provide the biosystematic backup to biodiversity.

Several special interest groups and networks (SIGN) dedicated to biodiversity conservation have sprouted up all over the global computer network. These SIGNs provide various facilities to users. Some of these are available to the email subscribers to the list. Subscribers usually join a list by sending an email with requisite information to the listserver or the master of the list. The user's name, email address and other information are then added to the email database maintained by the SIGN.

Biodiversity information networks in India: current status

There are a large number of Non-Government Organizations (NGO) and research organizations involved in conservation-related activities and research. India is not only rich in biodiversity, but can also boast of a large baseline information on this wealth. The repositories of this information are numerous and diverse, with some of the agencies being more than a century old. Invaluable information exists in a variety of forms: as specimens, field notes, reports and various kinds of computerized data.

Despite dramatic developments in information technology and quantum leaps in telecom facilities, the conservation efforts in India are yet to make the best use of the options available for information exchange, barring a few exceptions. The recent discussions on the mechanisms and scope of the information base for biodiversity conservation in India[1], is an indication of this realization. The information networks for biological conservation are at a formative stage, the front runner being the network initiative of the Foundation for Revitalization of Local Health Traditions (FRLHT).

The well-established and nationwide biological information network is the Bio-Technology Information System (BTIS) consisting of the Distributed Information Centres of the Department of Biotechnology supported by NICNET—the computer network of the National Informatics Centre (NIC). Another biological information service available is the access to the Medical Literature Analysis and Retrieval System (MEDLARS) over NICNET, provided to select institutions by NIC. The different institutions with access to MEDLARS do not, however, exist as inter-connected nodes.

Although FRLHT has taken an initiative in creating the Indian Medicinal Plants National Network of Distributed Databases (INMEDPLAN), it falls short of the larger needs of networking the conservation activity and bioresources mapping. This can only be met by the 'inter-networking' of several networks such as INMEDPLAN, BTIS and those in the offing like the Biodiversity Information System (BIS) of the Indira Gandhi Conservation Monitoring Centre (IGCMC) established by the World Wide Fund for Nature (WWF) and the Environmental Resources Information System (ERIS) being devised at Wildlife Institute of India (WII). The IGCMC proposals, unlike that of FRLHT, appear to favour an over-centralized system, rather than a true network of several fully autonomous, distributed information systems.

The INMEDPLAN is a network of organizations involved in identifying, documenting and reviving folk and traditional health systems and in the inventorying of medicinal plants. FRLHT has also attempted to bring about some sort of standardization in data compilation formats. A large number of organizations are also involved in the consultative process of standardization.

Although, the BTIS does not have conservation priorities, it provides rapid access to a wide variety of biological databases, besides those related to biotechnology. Moreover, considering the role that molecular biology and DNA fingerprinting are destined to play in biodiversity inventorying and monitoring, it is important to bring BTIS within the ambit of conservation network.

A recent study of biological databases available in India indicates that the majority of these are developed under BTIS[2]. Although the study does not appear to have examined the information resources available at reputed ecological research institutions, its overall conclusion that the contribution from sources other than BTIS is marginal, may have some validity. The same study also notes that the use of the available databases is restricted due to lack of electronic data transfer facility and networking. The BTIS users make use of NIC's

satellite-based network for data communication. However, BTIS does not function as a true on-line computer network due to certain limitations imposed by NICNET.

Both ERIS/WII and BIS/IGCMC/WWF are at an early developmental stage. It is important that the originators of these information systems lay down some long term networking and standardization goals as part of the design and development activity. What needs to be clearly recognized is that the highly centralized information systems relying on high performance computers are a thing of the past and the future belongs to inter-networking of special interest groups. With a highly decentralized and participatory approach to conservation, it becomes all the more important to anticipate the need for a truly extensive data network and take into account the multiplicity of end-user needs.

As important as access to information or even more, is the access to people and person-to-person communication on the network. It is such communication that makes the network come alive. The conservation effort in India, has an overwhelming need to create such a 'live' network, particularly because, informal networks of people and organizations exist and have been able to open up large areas of co-operative endeavour.

Information sharing: some contentious issues

The CBD introduces new elements into the ticklish question of information sharing across political boundaries. Even where national boundaries are not involved, there are many important matters of protocol and copyrights as well as the rights of local communities that need to be addressed for the network to be operational. However, it is better to take on the bull by its horns.

It has been argued, albeit with good intentions, that electronic exchange of biodiversity information will open a Pandora's box of possible infringements of national rights over genetic resources caused by information flow over global networks. Other contentious issues relate to intellectual property rights (IPRs) and cultural heritage rights in the context of CBD.

There is a view that until and unless these issues are fully resolved, it will be rather premature to think of electronic networks. However, what needs to be taken note of is that most of the information that may be available on these networks will in case be published sooner or later in journals. Publication of findings *per se* cannot be and should not be stopped, irrespective of the media used. Once published in journals, anyone around the globe can put that information out on a network. Thus, any blanket ban on information transfers on network cannot make sense today, when scientific journals are being available on CD-ROMs and almost every important journal is providing back issues as CD-ROM volumes.

It needs to be noted that in countries where the IPRs are enforced, rather stridently, that has not become a damper to information flow on networks. These services are being made use of in India, too, by scientists at research institutes in physical sciences and molecular biology as well as by the software developers. The IPR issues need not be a detriment to the emergence of CiNet. Besides, the network will only allow access to information that is open to public. It does not permit anyone to access any kind of information; all access being subject to the restrictions imposed by the network administrators.

The CiNet

The goal of CiNet is to create a voluntary, on-line 'live wire' network of researchers, government officials, ecosystem managers, executive decision makers, NGOs and local communities for biological conservation and sustainable development in India. It must be based on best available, cost-effective technologies and workable administrative arrangements. Instead of setting out to make large investments in infrastructure building, the network must make the best use of available facilities and function on a cost-sharing, no-profit-no-loss basis. To cut costs, the existing networks and access providers can be made active participants in this process.

The CiNet must play the following two roles: a) organize a loose network of organizations interested in conservation and b) establish a well-coordinated data network with features and functions similar to the special interest computer networks that exist elsewhere in the world.

The CiNet can succeed only by cooperating with other initiatives and by assuming an inter-networking role. The necessary administrative and technical protocols can be worked out to ensure meaningful cooperation and active collaboration among such organizations. The network must provide mechanisms at different levels for free information flow between these numerous organizations and between networks. The conceptual framework articulated here attempts to address the data flow challenges of the 'massive, decentralized' bioresources mapping exercise with the active participation of people, suggested by Madhav Gadgil[3]. He contemplates a 'nation-wide village level programme of mapping of natural and man-made habitats on the scale of a hectare or so'. The Panchayat Level Bio-Resources Mapping Programme, that he advances, involves teams of 'barefoot ecologists' under the guidance of taxonomic experts and are referred to here as Bio-Resources Mapping Teams (BMT).

The proposed CiNet consists of a network of Distributed Biodiversity Information Centres (DBIC) for data compilation, information processing and dissemi-

nation services assisted by several Field Studies Units (FSU) and BMTs. Selected institutions active in ecological research and willing to assist the CiNet activities could act as Associate Resource Centre (ARC) for the DBIC. These ARCs will also function as value adding nodes on the CiNet. The ARCs are expected to 'add value' by way of analysis and critical evaluation of the data as well as assist the BMTs. The data flow framework of the network is given in Figures 1 and 2.

The distributed information systems can be networked via land and satellite links making use of the telecommunication facilities and datalinks already available in India. The CiNet National Centre (CNC) can also provide decision support services to the Ministry of Environment & Forests, Government of India (MoEF) and other governmental agencies. The CNC will also be responsible for the overall co-ordination and technical support for the CiNet. In addition, the CNC will also carry out appropriate clearing house functions as the information travels across the network. The clearing house mechanisms will ensure reliability of the network and enforce quality control and validation checks on the data that is added to the central databases. The CNC must, therefore, have a technical secretariat for its clearing house functions. Ideally, the CNC could be located at a site where the best facilities and expertise in information technology as well as bioresources mapping, inventorying and monitoring are available.

The user community

The end users of the network are envisaged at different levels of decision making and management:

a) Macrolevel executive decision makers, national policy makers, planners, national institutions, national level NGOs such as WWF (India), inter-governmental organizations under United Nations, international development agencies and international institutions for the implementation of the CBD.

b) Decision makers, policy makers, managers, scientists, NGOs and individuals at the state, district and community levels.

c) Microlevel planners, community organizations and grass root level NGOs.

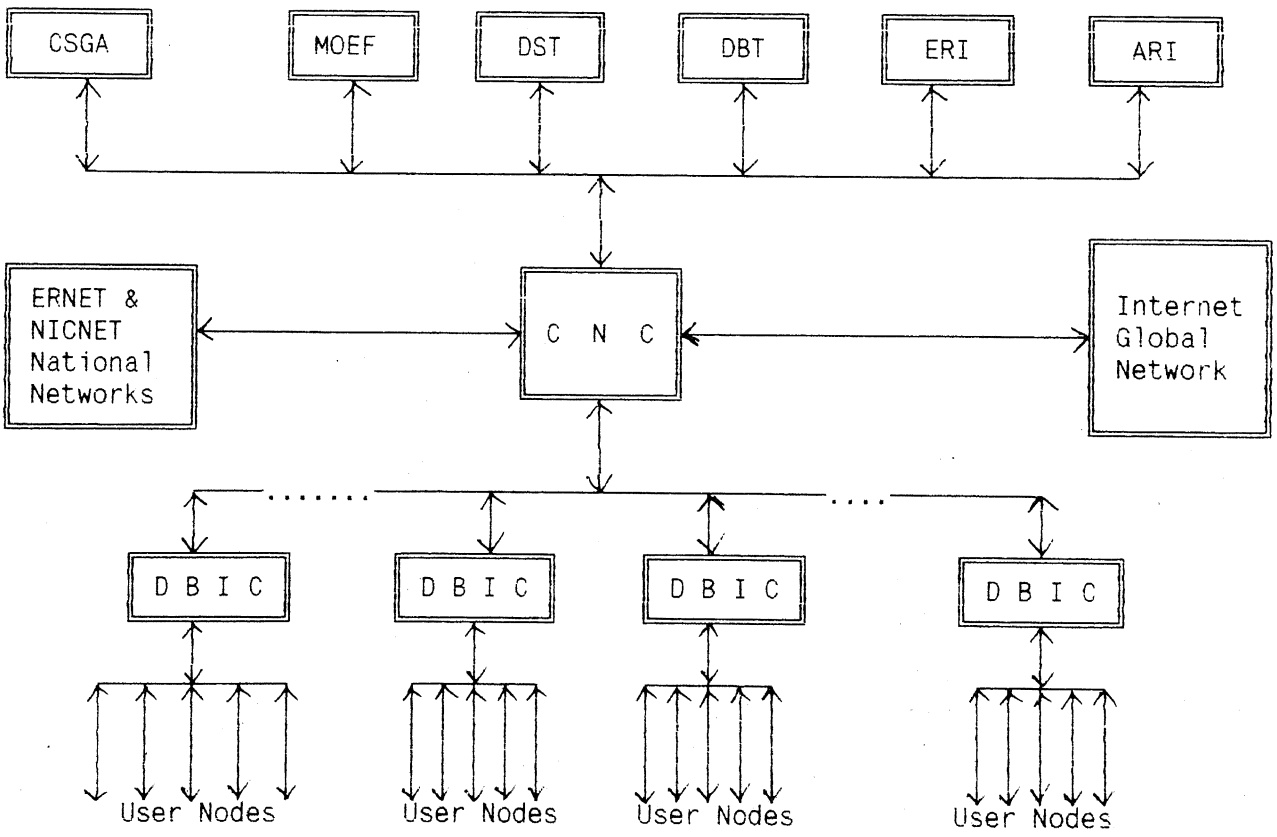


Figure 1. The logical CiNet with links to existing networks. ARI, Agricultural Research Institutions; CNC, CiNet National Centre; CSGA, Central and State Government Agencies; DBIC, Distributed Bioresources Information Centre of CiNet; DBT, Department of Biotechnology, Government of India; DST, Department of Science and Technology, Government of India; ERI, Educational and Research Institutions; MOEF, Ministry of Environment and Forests, Government of India; User Nodes, See Figure 2.

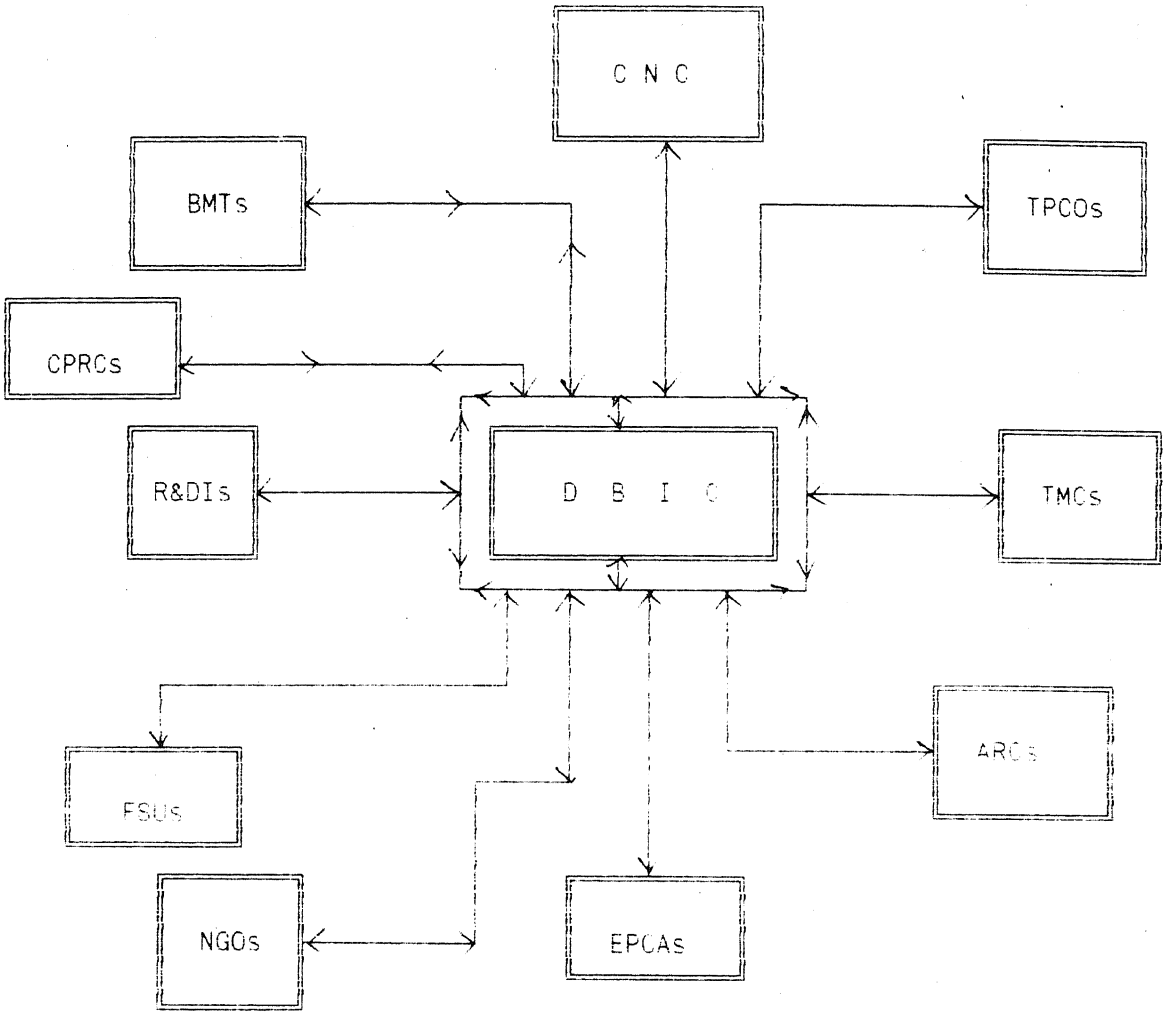


Figure 2. The logical network at a Distributed Bioresources Information Centre (DBIC). ARC, Associated Resource Centre; BMT, Bioresources Mapping Team; CNC, CiNet National Centre; CPRC, Common Property Resources Centre; DBIC, Distributed Information Centres of CiNet; EPCA, Environmental Protection/Conservation Authorities; FSU, Field Studies Unit; NGO, Non-Governmental (non-profit) Organizations; R&DI, R&D Institution; TPCO, Traditional People's Community Organization; TMC, Traditional Medicine Centre.

The CiNet can help the larger goals of biodiversity conservation by addressing the information needs of:

- protected area management
- short and long term ecological studies
- socio-economic studies and development programmes
- community-based development programmes
- adapting and improving the traditional methods of resource use
- local communities and bioresources mapping programmes.

The CiNet databases

Some of the major databases that must be organized under the CiNet are:

a) Data on the flora and fauna of India, with taxonomic details and information relevant for conservation of biodiversity.
b) Environmental impact and status data.
c) Sustainable resource use potential and practices; sustainable technological options; socio-economic profiles of sustainable resource use.
d) Hydrologic and agro-climatic and meteorological data.
e) Institutions, experts, resource persons and NGOs with information on area of specialization, infrastructure facilities, experience and catalogue information on educational software (printed materials and audio-visuals).
f) Bibliographic data extracted from on-line databases and those compiled from Indian publications not covered by international bibliographic services.
g) Metadatabases on information sources, specimen and

herbarium collections, map archives and photo/audio-visual libraries.

h) Socio-economic and demographic data relevant to ecosystem management and conservation.

Standards

The most important question that the network builders need to recognize and deal with is that of data standards. Certain standardization process is already on. However, a lot more remains to be done. Considering the large amount of information available in different forms, drawing up standards is an urgent need.

Some of the entities that must be accorded high priority for standardization (in alphabetical order) are:

- Abundance codes for trends and status.
- Bibliographic formats.
- Biodiversity mapping grid identification codes relating it to survey of India toposheets.
- Biogeographic site descriptors.
- Biosystematics.
- Conservation status descriptors for different levels of taxa.
- Degradation/disturbed area classification.
- Descriptors of anthropogenic pressures on ecosystems.
- Directory of castes and tribes based on the data from Anthropological Survey of India's Peoples of India Project.
- Distribution classes associated with different taxonomic levels.
- Ecological classification/habitat types and classes.
- Faunal and floral attributes or a simplified taxa identification key system.
- Landuse-related descriptors.
- Location identity/site codes.
- Macro- and micro-climate types.
- Resource use-related descriptors.
- Soil-related descriptors.
- Species–species relationships/associations.
- Water: Nutrient and water quality-related indices.

Geographical information system (GIS) and biodiversity maps

The CNC must be equipped with state-of-the-art work-station-based GIS facilities while at each DBIC, either a middle or entry level based GIS facility that can be run on a personal computer must be established or arrangements must be made for sharing GIS resources with a well-equipped ARC. The GIS makes it possible to make use of computers for the analysis of spatial data. The spatial data such as maps are first converted into computer readable form. This is achieved using hardware such as scanner or digitizer. GIS can also make use of satellite imageries and compare such images taken at different times. It is also possible to overlay these computerized maps with information collected from ecological surveys by the FSUs. The data residing on well-structured databases can also be incorporated into the GIS to produce spatial representation and visualization of such data. It is also possible to undertake substantive statistical analysis of the spatially aggregated data.

The GIS facilities at the DBIC must provide facilities for updates and retrievals of spatial data for use by the FSUs. Therefore, besides computer systems, requisite input/output and display devices like large high resolution colour video display units, digitizers, scanners and colour printers need to be installed at the DBICs. These facilities only need to be added to the DBICs in a phased manner keeping in step with the growth of the bioresources mapping programme.

CiNet needs to work towards developing a comprehensive GIS-based bioresources map library and make it available on-line to the CiNet users. Since the spatial scales envisaged for the biodiversity map grids under the bioresources mapping programme is of the order of 1 to 10 hectare grids, the map libraries must organize spatial data on maps of very large scale. In practice, it will be difficult to organize map libraries for 1 hectare grids and it may be necessary to opt for larger biodiversity mapping grids of nearly 1000 hectares.

Information-sharing mechanisms

The major databases of CiNet must be organized using database management system technologies that conform to industry standards. It is desirable that the large central databases of CiNet are well-structured based on relational database management concepts. Such well-structured databases can be used to hold the authenticated baseline data, with a bare minimum of essential details. It is also necessary to develop multimedia databases that can include text, pictures and sound. These databases will reside on servers providing access to data in several ways. However, as pointed out earlier, information on computers is not organized merely as well-structured databases; there is certainly a lot more to information sharing via networks, than databases. Besides well-structured databases, information needs of the network users have to be met in several other ways as well. A recent workshop that examined the linking mechanisms for BIN21 concluded that the best approach was for it to function as a SIGN having several participating nodes with web servers[4].

Typically, the information on a network will consist of non-database entities such as documents, graphic objects, messages, news bulletins, reports on new findings, electronic journals, file archives, and large amounts

of information organized in a hypertext format. The network has to provide nodes with servers that can provide a host of services available on the Internet, which is a 'network of networks' linking computers all over the globe. The CiNet can be built by creating an exclusive network of members or by becoming a part of the global Internet community. The later option is cost effective and it is possible to make use of the Internet connectivity in collaboration with the Education and Research Network (ERNET) project. The ERNET infrastructure can be augmented to meet the needs of CiNet. It must be noted that the users will have varying degrees of difficulty in network access. Therefore, a host of linking mechanisms, that allow connectivity by means of low speed dialup telephone lines as well as high speed links over leased telephone lines, satellite links and public data networks, must be made available to the CiNet users.

The networks on Internet maintain mail lists, news bulletin boards, special interest groups, file archives and help individuals find the right information or the right contact. The main Internet tools are Gopher, World Wide Web (WWW) and the Wide Area Information Server (WAIS). The information available on Gopher servers are limited to text. Gopher organizes information in a hierarchical menu-based system, analogous to the table of contents of a book. The WWW uses hypertext links to connect together pieces of information contained in documents located at different nodes. The information in the documents is organized using the HyperText Markup Language (html). Files are transferred using the HyperText Transfer Protocol (http) and the File Transfer Protocol (ftp). The WAIS is a client-server tool that can be used to search and retrieve files, based on full test indexing of the contents or titles.

These Internet services can be organized on a cost-sharing basis with the major initiative coming from a large organization or university. The network must provide all or some of these facilities, besides off-line email. The CNC and each DBIC must set up a WWW or Gopher. The CNC need to operate listservers and bulletin boards. The CNC also need to provide access to major databases. The DBIC must provide user access to on-line searches on the databases located at any of the DBIC or CNC.

It is important that besides email, CiNet users have access to Internet services such as ftpmail and mail-gopher. The DOS or Windows-based PCs connected to the DBICs using modems can be provided shell accounts on the servers at the DBIC. Those nodes which can be connected to the DBIC servers using high speed links or leased telephone lines can make use of the services such as Gopher, WWW or WAIS. Each of the BMT, FSU or ARC must have *at least* email access, and if possible other facilities based on TCP/IP. The TCP/IP (Transmission Control Protocol/Internet Protocol) is a way of packaging data for easy movement between many different computer systems. It consists of a collection of search, retrieval and communications utilities. Personal computers running DOS/Windows can make use of the network services by means of either the Serial Line Internet Protocol (SLIP) or the Point-to-Point Protocol (PPP) with the TCP/IP stack up and running on the PC, provided the DBICs offer SLIP or PPP accounts on the servers.

To facilitate network connectivity there are certain software and hardware needs. To enable simple email facility for a CiNet user equipped with a PC, all that is needed is a mail service account on the DBIC server, a telephone, a modem and a communication software loaded on the PC to enable terminal emulation after getting connected to the DBIC server. The DBICs, also need to have facilities for remote users to login via telephone lines and must therefore have the DBIC server also connected to the telephone lines through modem. The modem speeds in this can be even as low as 300 bits per second (bps) and faster modems can be used where higher speeds can be attained. The computer at the user end can also be a multi-user system running UNIX. If only a dialup link with the DBIC is possible, then an UUCP (Unix-to-Unix Copy) account can be created on the DBIC server to transfer email as a fully automated service.

If the user has access to better communication facilities such as the X.25 based connection or a leased high speed telephone line to a DBIC server, and the computer system at the user end is multiuser running any port of System V Release 4 UNIX complete with X/Windows Graphical User Interface and TCP/IP, then all the network services can be made use of. Email can be transferred using the Simple Mail Transfer Protocol (SMTP) rather than UUCP. Several other facilities such as FTP and TELNET for remote login to the CiNet servers will also become possible. A high speed modem is needed because if the user is to make the best use of network services, connectivity must be achieved at speeds of 14.4 kbps or above. Otherwise the network access can become very tiresome. High speed modems and interface cards are required at both ends.

Software support is needed by way of the communication protocols (TCP/IP and SLIP or PPP). Both the PC at the user's end of the CiNet and the DBIC server of the CiNet must be running either SLIP or PPP as well as the versatile TCP/IP. With the software and hardware (interface cards on the computers and the modems) in place, the user will have full access to all the public information resources on the CiNet as well as Internet, assuming that CiNet has full Internet connectivity.

The field units which are the key to the biodiversity mapping programmes also need to remain 'online' with

the DBICs. Very often the FSUs may be conducting surveys in remote locations with no possibility of access to telephone links. It is, indeed, possible to resort to wireless data transmission to access the nearest DBIC when such a need arise. There are, of course, certain legalities to be sorted out in order to realize this technical possibility because of the restrictions imposed by the laws governing wireless communication. Sometime in the near future, mobile telephones that can be used from any geographical location will be available. The costs of such systems may, indeed, be rather high at the introductory stage. When the use of such devices becomes more widespread and cheap, the FSUs can use it for establishing dialup links with the DBICs.

Communication options

The CiNet is conceived as a loose network of organizations that can share information on a cost-sharing basis. It can make use of the different computer networks available such as the satellite-based NICNET services using Very Small Aperture Terminals (VSAT) and the ERNET services based primarily on terrestrial dial up links. ERNET, too, is expected to provide access via VSAT soon.

There is no need to go through the whole cycle of developing networks and physical links, since the network can be developed using the data communication facilities available in the public and private sectors. In a national context where many data networks already exist, the feasible low investment proposition is to make the best use of existing network infrastructures on a cost-sharing basis and let the CiNet 'evolve' on its own as the activity increases. ERNET with its vast experience in networking, can possibly play a pivotal role in establishing the required network facilities and CiNet can ride over the ERNET.

There are, indeed, several communication options available. Depending on what is appropriate at each level of the data flow, the best option has to be exercised. The national network of DBICs and the CNC that make up the CiNet need to be linked by high speed links. The VSAT based links is a reliable option. Another possible option is to go in for the INET facilities offered by the Department of Telecommunications (DOT) which allows X.25 based communications. The X.25 is the name for a widely used network layer protocol described by the International Standards Organization (ISO) and the International Telegraph and Telephone Consultative Committee (CCITT).

The wide area network (WAN) services can be implemented by establishing permanent physical connections using dedicated lines or private lines, by circuit-switched physical connections based on dialup links and by demand digital service based on packet switching. A network of networks, typically, uses packet-switching methods. Rather than trying to establish a dedicated communication line between two computers located in the same or different network, which is rather clumsy when several computers are competing for connections, the inter-networking typically uses packet switching techniques. This is because packet-switching allows several computers to share communication links, by transmitting packets of information with an addressing system through the packet switching circuit.

Packet-switching has in-built error detection and correction mechanisms and provides dynamic re-routing of calls as well as interconnection of computers/terminals at different speeds. It will be far too expensive to build a private data network dedicated only to CiNet activity. The facilities provided by DOT on INET—a X.25 based Packet Switched Public Data Network—include permanent point-to-point connection between two end users of the circuit or multipoint connections between more than two end users.

Given the communication options available, the pros and cons need to be weighed carefully. While making the choice the following aspects are to be considered:

a) The communication option must ensure high speed connectivity to major national and international networks, particularly the NICNET, ERNET and INTERNET.
b) The communication link should not be prone to frequent breakdowns.
c) The recurring costs must be kept low.
d) The communication facility must support multiple protocols.

As network service providers, particularly in the private sector, make their bid for a slice of the Indian cyberspace, more options will be available for networking and CiNet itself can possibly become a value-adding, financially self-sustaining service provider with its specialized information resources.

Conclusion

The conservation of biodiversity is an enormously complicated and difficult task requiring information flow on an unprecedented scale. The conservation efforts need networking of individuals, organizations and information systems. Co-operation and collaboration on such a wide scale will be possible only with the support of an appropriate information network, such as the one presented here. The user demand, technical expertise and communication infrastructure for the creation of such a network exist in India and the network proposed here is envisaged as a self-sustaining one that is well

integrated into the existing data networks, particularly the rapidly expanding Internet community.

1. Pillai, J., *Curr. Sci.*, 1994, **66**, 892–893.
2. Chavan, V. and Chandramohan, D., *Curr. Sci.*, 1995, **68**, 273–279.
3. Gadgil, M., *Curr. Sci.*, 1994, **66**, 401–406.
4. Canhos, D. A. L. (ed.), Proceedings of a Workshop for the Biodiversity Information Network, Base de Dados Tropical, Campinas, Sao Paulo, Brazil, 1994.

ACKNOWLEDGEMENTS. This work draws inspiration from the large scale bioresources mapping programme articulated by Prof. Madhav Gadgil. I thank the referee K. Gopinath, Indian Institute of Science, Bangalore, for the critical comments and Gopi K. Garge, Network Coordinator, ERNET, Indian Institute of Science, Bangalore who contributed valuable suggestions and ideas. This paper is dedicated to my late colleague Dr D. F. Singh, who was heading the Wetland Ecology Division and was working towards creating a database on fish diversity of the Western Ghats on CD-ROM.

Earth's core

S. K. Saxena

The new techniques of heating solids with a laser at ultra high pressure in a diamond-anvil cell are providing us with new data on equation of state properties of solid and molten iron. The new data on iron melting, which extends to a pressure of nearly 2 megabar (200 GPa), permits the thermodynamic extrapolation of the melting data to pressures corresponding to the centre of the earth. The new studies on iron have also prompted a search for new phases of iron possible at high pressure and temperature. The currently available data point to a rather simple compositional model of the core which is identical to the meteorite irons or stony-irons. The adiabatic temperature of the earth's centre may lie between 5980 and 6680 K.

SEISMIC data show that the earth's core consists of an outer part which has all the properties of a liquid and an inner core which is solid (Figure 1). The core, reaching to a depth of nearly 2900 km from its centre stores a substantial part of the planet's energy and, therefore, exercises significant influence on the dynamic processes in the mantle. From cosmochemical, geophysical and geochemical considerations, the dominant species in the core has to be iron. Iron occurs in four solid structural states, namely, δ (bcc), α (bcc), γ (fcc) and ε (hcp). The ε (hcp) is considered to constitute the bulk of the solid inner core. Although not experimentally shown but based on the available data on the compressibility of iron (ε), it appears that iron by itself has a density higher than the density of the core[1]. Therefore, it has been suggested by many geochemists that some additional light elements (e.g. sulphur, oxygen, Si) are mixed with iron to form a lower density inner core and a lower melting liquid outer core[2].

To understand the state of the core, it is important that the thermodynamic properties of liquid and solid iron are determined. Recent developments in generating ultra-high pressures in diamond-anvil cells with laser heating have made it possible to study properties of material at the core pressures and temperatures. The purpose of this article is to report these technical developments, the recent findings on iron and the status of our understanding of the earth's core.

Figure 2 shows a summary of the currently available experimental data, until the year 1993, on phase equilibrium relations obtained from static devices, i.e. techniques with *in-situ* heating of a sample under pressure

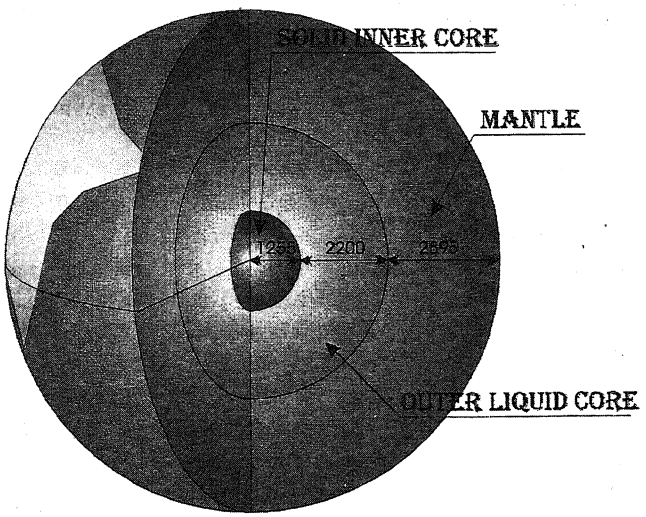


Figure 1. Earth's interior as determined by seismological data. The solid inner core with a radius of 1255 km is surrounded by a layer of molten iron with a shell thickness of ~ 2200 km reaching a depth of nearly 2900 km from the crust. The inner core density is close to 13 g cm^{-3} and its temperature is variously estimated as between 4000 and 8000 K.

S. K. Saxena is in the Institute of Earth Sciences, Uppsala University, S-75236, Uppsala, Sweden.

sustained over a period of time. The phase transition boundaries α–γ, γ–ε and α–ε meet at one point, the so-called triple point; at such pressure and temperature all the three forms of iron coexist in equilibrium. Recent theoretical study on iron properties at high pressure and high temperature by Anderson[3] has shown that an experimental determination of the data on ε–γ phase transformation and the ε–γ-melt triple point are particularly crucial in extending our currently available results to understand the state of the earth's core. The argument is that the pressure-temperature of the triple point determines the behaviour of the melting curve at high pressures. Thermodynamically, the pressure-temperature slope of the ε–γ phase equilibrium curve constrains the estimated enthalpy, entropy and other physical properties of the ε phase which are largely unknown. With the availability of the equilibrium data on the ε–γ phase transformation, one may assess the thermodynamic properties of the ε phase which in turn can be used in the estimation of the properties of the iron melt.

A triple point, similar to the one shown in Figure 2, should exist for the coexisting γ, ε and melt. All the available experimental data on the ε–γ transition (see Figure 2), appear to indicate that such a triple point would lie around 2700 K between 60 and 70 GPa. The problem is that Brown and McQueen[4] had located a solid–solid transition some years ago at 4400 ± 300 K at a pressure of 200 ± 2 GPa. As shown in Figure 2, this phase transition could not be the ε–γ transition. The triple point is located at too low a pressure for this to be possible. This apparent inconsistency between the sets of data led Boehler[5] to conclude that there should exist a fifth iron phase. The existence of this phase is supposed to solve two problems; first it would explain the phase transformation recorded by Brown and McQueen in their shock-wave studies and second, by creating a third triple point, it would let the melting curve slope appropriately to fit the existing geophysical data. This led to the start of a search for a new phase of iron and the determination of the iron melting at pressures higher than that at the core/mantle boundary (130 GPa). Achieving such extreme pressures in the laboratory requires special techniques, for example, the laser heated diamond-anvil cell.

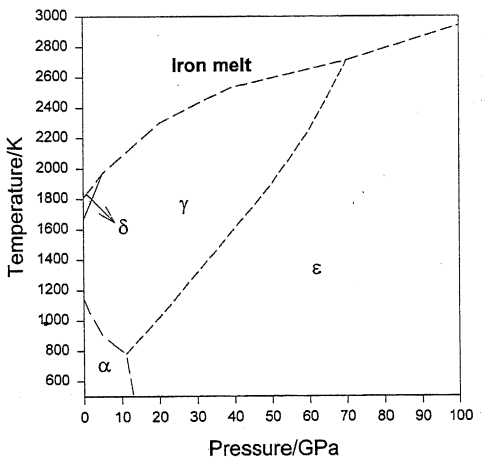


Figure 2. Phase diagram of iron from the available data up to 1993. The various curves drawn do not include the new experimental data. The stability regions are shown for the four iron polymorphs: δ (bcc), α (bcc), γ (fcc) and ε (hcp).

The diamond-anvil cell experimental set up for generating high pressure is discussed in several recent articles[6]. Here a brief description specific to Uppsala Laboratory is given (Figure 3). The laser heating system most useful for such studies is the one developed by Boehler *et al.*[7]. The use of a high power Nd : YAG laser (e.g. the one by Coherent) operating in CW TEMoo mode at a wavelength of 1064 nm is required. The laser provides maximally 35 W of vertically polarized light with a feedback controlled stability better than 0.5% peak-to-peak. The high pressure requires the megabar diamond anvil cell developed by Jephcoat *et al.*[8] for such experiments. To achieve real high pressures beyond 100 GPa, one needs to use the standard Drukker type IA beveled diamond anvils with 0.30 mm diameter flat culet faces (Figure 4). For the pressure determination the calibrated pressure shift of ruby R1 fluorescent line is used[8]. A phase transformation is determined by visual observation of melting when possible and by plotting the laser power against the temperature. At the temperature of phase transformation there is a distinct change in slope. The temperature is determined by spectroradiometry. The 50 μm in diameter slit of the

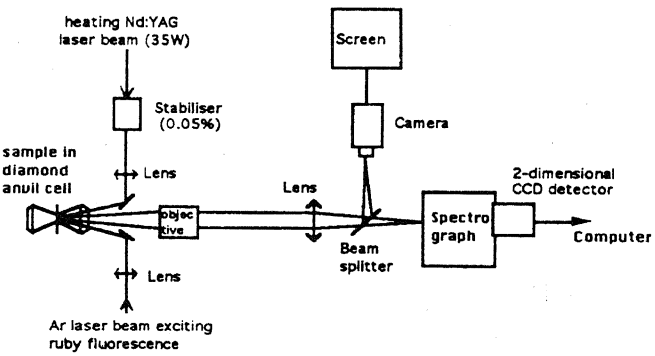


Figure 3. Experimental set up for the study of iron in a laser-heated diamond-anvil cell.

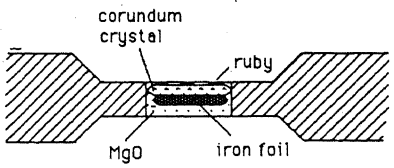


Figure 4. The sample environment in the experiment on melting. The iron foil is cushioned with MgO powder with a corundum plate on the top. Ruby chips are used for pressure calibration.

spectrograph collects the radiation from an area of 3 μm in width of the sample. A two-dimensional CCD detector is used to obtain temperatures measured over the entire length (30 μm) of the slit area. The temperature gradient over the central area (3 μm) is only a few degrees. The laser power fluctuations are minimized to a root mean square value < 0.5% by using a laser power stabilizer. In the method suggested by Boehler *et al.*[7] a ruby crystal is used on the top of the sample as an isolator and MgO powder at the bottom. The use of crystals eliminates the possibility of any significant reaction between the sample. In particular, at the sub-solidus temperature range of such study, reaction between the iron and the ruby is not a problem. An area of 3 μm^2 is sampled by the spectrograph to get one pressure value; this reduces errors in pressure determination resulting from the presence of pressure gradient in the chamber. The pressure is determined at room temperature. Following the discussion by Heinz[9], a 7% decrease in the cold pressure may be made to account for the thermal pressure. Prior to measurement, a chosen spot is heated by laser to mechanically relax sample assemblage. The pressure usually drops up to 16% of the initial unrelaxed value. The pressure drop is almost linear and can be stabilized by repeated heating. The following problems are cited with the diamond-anvil cell experiments: i) reaction with the pressure medium and ii) temperature and pressure gradients. We dry the cell carefully. The corundum plate in contact with the iron does not show any visible reaction. Data on melting at high pressures show that corundum melts at higher temperature than iron at all pressures studied. We measure temperature over an area of less than 3 μm where melting was observed directly at lower pressures. It is further ascertained by the change in slope on a plot of the laser power against temperature. At high pressures (above 40 GPa) visual observations become difficult and then one must use the laser-power/temperature function and quench textures. Saxena *et al.*[10] demonstrated the accuracy of the temperature measurements by melting several solids with known melting temperatures.

In Figure 5 results of some recent experiments are displayed. The figure shows the stability of a new phase (called β by Saxena *et al.*[10]) and melting of iron to a pressure of 150 GPa by Saxena *et al.*[10]. Melting data are available to a pressure of nearly 200 GPa by Boehler[5]. The issue of the new phase is not fully resolved. It is likely that we have many additional structural modifications of iron. To begin with it is possible that the phase transformation reported by Boehler and Saxena *et al.* are for two different phases. In a recent study, by using the laser heated facilities at the Brookhaven National Laboratory synchrotron X-ray facilities, Saxena *et al.*[11] found that the hcp-iron (ε) transforms to a dhcp-iron (a four-layered hcp structure) at about 35 to 40 GPa at approximately 1500 K. This is certainly a confirmation of the possibility that we may have one or more yet unknown polymorphs of iron to be discovered.

Saxena *et al.*[10] used the new experimental data to obtain an internally consistent thermodynamic database with which to model the iron phase diagram and the core energetics. The newly generated data may then be used to calculate iron phase diagram. This combination of thermochemical and experimental data provides us with an estimation of temperature gradient in the earth's core.

Figure 6 shows the calculated phase transformations and melting in iron at pressures from 0 to 350 GPa. The extrapolation of the melting curve is done by calculating the melt compressibility as a function of temperature from the experimental data of Boehler and from this study (model 1). In such assessment of thermodynamic data, we did not include the nearly 10 to 15% error of pressure measurement possible when pressure exceeds 100 GPa (ref. 8). We generated a second model by permitting the melt compressibility to decrease at a higher rate at high temperatures by including the lower limits of the shock-wave data from Brown and McQueen[4] (see Figure 6). The high temperature melt data should be shifted to lower pressures by approximately 10%, which is quite within the stipulated random errors in the pressure measurements from the ruby fluorescence[8]. On such a melting curve (model 2), there are two triple points; the first (melt-β–γ) is located at 2946 (± 100) K and 76.5 (± 4) GPa; the second is located at 4580 K and 216 GPa insuring that the shock-wave constraint on melting[4] is satisfied. If we use the

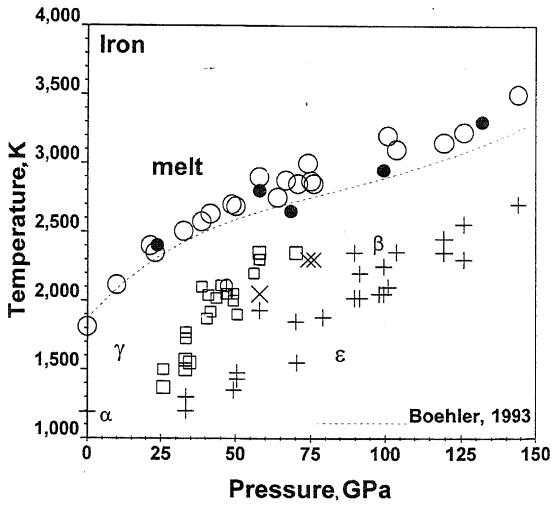


Figure 5. The experimental data on iron phase transformation and melting generated at Uppsala. The open circles represent the melting temperature determined from either visual observation or using the laser-power/temperature slope change. The size of the circles represents the error in temperature due to laser fluctuation and statistical errors of fit. The solid circles represent that iron remained solid up to that temperature. The squares represent the e–g transition and the plus sign the new phase β–ε transition. Crosses represent a phase transition which could not be labelled as either the β–γ or ε–γ.

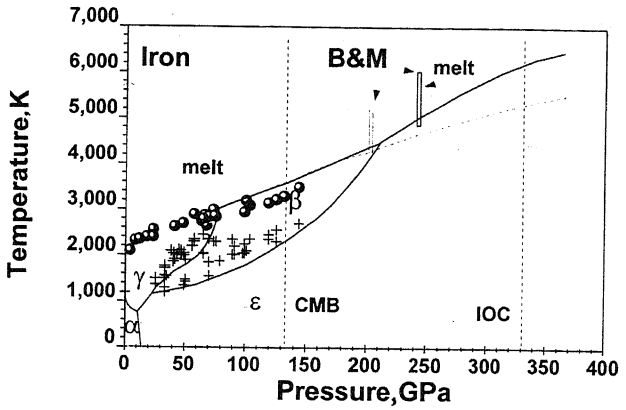


Figure 6. Iron phase diagram as calculated from the thermodynamic data from Saxena *et al.*[11]. The dashed line represents an extrapolation of the melting curve (thermodynamically calculated) with no consideration of the data by Brown and McQueen[4] (labelled as B & M). The solid line represents the preferred model and includes the consideration of the latter data within the experimental errors of pressure determination.

low-temperature limits of these data as found through model 1 which gives the second triple-point (melt-β–ε) at 4015 K and 205 GPa.

A model of the earth's core should (a) explain why the inner core is solid and the outer core liquid, (b) must use a chemical composition which is consistent with cosmochemical data and (c) be similar in density to the seismic density. We note that the answer to the first problem lies in the special melting characteristics of iron associated with phase transformations and the location of the triple point ε–β-melt. The calculated melting curves in Figure 6 show that we may expect iron melting to occur at central core pressures (363.85 GPa, PREM model[12]) between 5980 (model 1) and 6680 K (model 2). Then if we assume that the temperature in the inner core is largely adiabatic, we may start the calculation of the adiabat starting where the pure iron melting curve intersects the inner-core/outer-core seismic discontinuity. The calculated temperature (model 2) at the earth's centre is 6150 K and at the top of the inner core 6130 K. The calculated densities of ε-iron just below the melting temperatures (13.76 g mol^{-1} at 5979 K and 13.55 g mol^{-1} at 6679 K) are considerably higher than the PREM model densities. Similar conclusions led Jephcoat and Olsen[1] to suggest that the inner-core must contain a few per cent of a light component such as pyrite. A light component both reduces the density and the iron-melting temperature. To maintain a solid inner core with the appropriate PREM densities, it is sufficient to add a small amount of Mg_2SiO_4 (high pressure equivalent phases: perovskite + periclase) to ε-iron (Figure 6). Both perovskite and periclase are refractory phases, with melting temperature higher than iron[13]. At the centre, the correct density is produced when we add 1.05% solids at 6150 K. In the inner core, it was not necessary to change the composition to obtain correct densities at slightly decreasing temperatures until the geotherm meets the inner/outer core boundary at 6133 K close to iron melting at 328.85 GPa. For the geothermal gradient to change smoothly and approach the outer core and mantle boundary at ~ 4000 K, the mole per cent of solids should increase from 0.7 at the bottom of the liquid core to 1.64 at the top. For a given composition, the outer core densities do not change significantly with temperature. However, a small decrease in the amount of added solid raises the temperature significantly. With model 1, we found the inner-core temperature to vary from 5430 to 5419 K and PREM density is matched with 1.5% of solids. A little additional amount of solids (~ 1.8%) is required to match the density in the outer-core along the thermal gradient shown in Figure 6. Although Anderson[14] used different thermodynamic arguments, his calculated temperatures of 6450 K for the centre and 6210 ± 400 K for the inner and outer core boundary are quite similar to the temperatures shown here. The second model is also in general agreement with the results of Brown and McQueen[4].

The experimental iron phase diagram and the thermodynamic data result in a rather simple model for the core. The relationship between the earth and meteorites has been discussed thoroughly in the literature. The class of meteorites called irons and stony-irons are considered to be debris from differentiated broken up planetary bodies. What the present results tell us is that this picture fits the model of the earth's core very well. Thus the core has simply the same composition as the irons and stony-irons. The density from such compositions fits the seismic density well. Dilution of iron by addition of sulphur or another light element is not essential. The outer core is liquid because of the phase transformations in iron and its melting behaviour.

1. Jephcoat, A. and Olson, P., *Nature*, 1987, **325**, 332.
2. Anderson, D. L., *Theory of the Earth*, Blackwell Scientific Publications, Boston, 1989.
3. Anderson, O., American Institute of Physics Conference Proceedings, Colorado Springs, Colorado, 28 June–2 July, 1993.
4. Brown, J. M. and McQueen, R. G., *J. Geophys. Res.*, 1986, **91**, 7485–7494.
5. Boehler, R., *Nature*, 1993, **363**, 534–536.
6. Saxena, S. K., Shen, G. and Lazor, P., *Science*, 1993, **260**, 1312.
7. Boehler, R., von Bargen, N. and Chopelas, A., *J. Geophys. Res.*, 1990, **95**, 21731–21736.
8. Jephcoat, A. P., Mao, H. K. and Bell, P. M., *Hydrothermal Experimental Techniques* (eds Ulmer, G. C. and Barnes, H. L.), Wiley Interscience, New York, 1987, pp. 469–506.
9. Heinz, D. L., *Geophys. Res. Lett.*, 1989, **17**, 1161–1164.
10. Saxena, S. K., Shen, G. and Lazor, P., *Science*, 1993, **260**, 1312.
11. Saxena, S. K., Shen, G. and Lazor, P., *Science*, 1994, **264**, 405–407.
12. Saxena, S. K., Dubrovinsky, L. S., Häggkvist, P., Cerenius, Y., Shen, G. and Mao, H. K., *Science*, 1995, in press.
13. Zerr, A. and Boehler, R., *Science*, 1993, **262**, 553.
14. Anderson, O. L., *Geophys. J. R. Astron. Soc.*, 1986, **84**, 561.

RESEARCH ARTICLE

Effect of micellar and microemulsion media on the oxidation of iodide by peroxodisulphate reactions

J. Santhanalakshmi and R. Kalaivani

Department of Physical Chemistry, University of Madras, A.C. College Campus, Madras 600 025, India

Water–sodium dodecylsulphate (SDS)–oil–*n*-alkanol microemulsions are constructed and used as media for I^- oxidation by $S_2O_8^{2-}$ ions. Normal SDS micelles and microemulsions in aqueous medium catalyse the self-decomposition of $S_2O_8^{2-}$ and $S_2O_8^{2-}$–I^- oxidation reactions. The overall and individual rate constants are obtained. Temperature, cosurfactant and oil effects are studied. Michaelis–Menten-type kinetics involving micelle–substrate complex equilibrium seem to describe the oxidation reaction. The E_a and ΔH variations depict the interfacial effect on substrate binding.

MICROEMULSIONS (ME) belong to a special type of surfactant assemblies in solution containing nearly monodisperse micelles or oil solubilized micellar droplets, with sizes ranging from 10 to 100 Å (refs 1,2). ME also serve as potential media for promoting compartmentalized, enzymatic, catalytic and encapsulation reactions[3,4]. In conventional ME-catalysed reactions, the micro-environments inside the droplets serve as microreactors in molecular scale for reactants solubilized in the oil or micellar cores[5,6]. Otherwise, when the dispersion medium of ME or micellar solutions serve as the reaction medium, the rate accelerations may be attributed to the presence of microdroplet or micelle spheres which create interfacial changes, excess interfacial area for substrate binding, decrease in the approachability barrier for similarly charged reactants and creation of parallel reaction of the same kind from the microemulsions and micellar surfaces, etc.[7]. Through suitable control of size of droplets, rate accelerations may be controlled and achieved through variations in the compositions of ME components, i.e. oil, water, surfactant and cosurfactant[8,9]. The oxidation reaction between iodide (I^-) and persulphate ($S_2O_8^{2-}$) ions has been chosen to monitor the ME and micellar effects because the kinetics of this reaction in aqueous medium with pH, salt and medium effects are well studied[10,11]. Since anionic reactants are involved, presence of cationic surfactants may catalyse the reaction. However, when rate accelerations are noted for anionic surfactant containing micellar and ME solutions, then ME and micellar media exemplify the concepts of micellar and ME catalysis and highlight the creation of appropriate interfaces, substrate binding, stability factors, etc.[12]. The role of microenvironments of water droplets in water in oil (W/O) aerosol OT ME on the I^- oxidation by $S_2O_8^{2-}$ has been studied[6]. Regarding oil in water (O/W) ME, however, no such study prevails. Therefore, we report here the utilization of O/W ME-containing sodium dodecylsulphate (SDS), oil, water, *n*-alkanol (cosurfactant) for enhanced rate of oxidation of I^- by $S_2O_8^{2-}$.

As cosurfactants and oil in nature form the essential parts of ME in creating the interfacial area and oil microdroplets, the effects of cosurfactant (isopropanol, *n*-propanol, *n*-butanol) and oil (benzene, toluene, xylene) are studied along with salt effect (Na_2SO_4) and activation-energy studies. Substrate binding constants on micellar and ME droplet surfaces, rate constants for $S_2O_8^{2-}$ self-decomposition and I^- oxidation are determined through kinetic principles derived in a simpler manner.

Theory

The presence of peroxo (O–O) linkage in $S_2O_8^{2-}$ anion induces easy self-decomposition to appreciable extents even under normal temperature conditions. In oxidation reactions (of I^-) involving $S_2O_8^{2-}$ anions at $T > 25°C$, fractionation of $S_2O_8^{2-}$ composition into parallel reactions, i.e. (k_d) self-decomposition and substrate oxidation (k_o) and with the separation of individual rate constants from the overall rate constant need to be analysed and when T is low, $k_d < k_o$, while T is high $k_d \geq k_o$. In presence of SDS and ME, both k_d and k_o are accelerated. k_d and k_o represent first and second order reaction rate constants. The normal mode of I^- oxidation may be considered to follow a Michaelis–Menten-type scheme since initial iodoperoxodisulphate [$IS_2O_8^{3-}$] intermediate formation is feasible due to the higher oxidation potential of $S_2O_8^{2-}$ than I^- and subsequently leads to product formation[13].

$$\text{I. } S_2O_8^{2-} \xrightarrow{k_d} 2SO_4^{2-}.$$

$$\text{II. } S_2O_8^{2-} + 2I^- \underset{}{\overset{K_o}{\rightleftharpoons}} I^- + IS_2O_8^{3-} + M \overset{K_m}{\rightleftharpoons} MIS_2O_8^{3-} + I^-$$

$$\text{(a) } \downarrow k_{20} \qquad \text{(b) } \downarrow k_{2m}$$

$$P \qquad P + M$$

where $P = 2SO_4^{2-} + I_2$; M = SDS micelle or ME droplet.

The overall kinetics of steps I and II are monitored through I_2 formation with respect to time and under pseudoconcentrations with $[I^-] = 10\,[S_2O_8^{2-}]$. Rate $= dx/dt = d[I_2]/dt = k'\,[S_2O_8^{2-}]\,[I^-]$ under pseudoconditions,

$$k'[I^-] = k;\ \text{rate} = k\,[S_2O_8^{2-}]. \tag{1}$$

Step I is first-order. Therefore at any time t, $[S_2O_8^{2-}]_t = [S_2O_8^{2-}]_o\, e^{-k\,dt}$. Hence equation (1) is

$$d\,[I_2]/dt = k\,[S_2O_8^{2-}]_o\, e^{-k\,dt}.$$

Integration between $t = 0$, $[I_2] = 0$ and $t = t$; $[I_2]$ limits gives

$$[I_2]/k\,[S_2O_8^{2-}]_o = (1 - e^{-k\,dt})/k_d$$

$$= (e^{k\,dt} - 1)/k_d\, e^{k\,dt}. \tag{2}$$

Expansion of $e^{k\,dt}$ series, neglecting t^2 powers, etc., the RHS of equation (2) becomes

$$k\,[S_2O_8^{2-}]_o/[I_2] = k'[I^-]_o\,[S_2O_8^{2-}]_o/[I_2] = 1/t + k_d. \tag{3}$$

Thus, applying linear regression on $[I^-]_o[S_2O_8^{2-}]_o/[I_2]$ against $1/t$, values of $k = 1/\text{slope}$ and $k_d = k'x$ intercept are obtained. Likewise, k_d in presence and absence of SDS, and ME may be derived. From known k_d values, fraction $[S_2O_8^{2-}]$ decomposed in t is known and the remaining fraction of $[S_2O_8^{2-}]$ contributes to the oxidation reaction. In normal reaction, the overall rate constant of oxidation reaction (k') may be considered as k'_o, and as k'_m in presence of SDS and ME. Thus second order rate of II(a) reaction will be

$$\text{vII(a)} = k'_o\,[S_2O_8^{2-}]\,[I^-]. \tag{4}$$

In terms of the reaction scheme,

$$\text{vII(a)} = k_{20}\,[I^-]\,[IS_2O_8^{3-}] \tag{5}$$

$$= \frac{k_{20}\,K_o\,[I^-]^2\,[S_2O_8^{2-}]}{1 + K_o\,[I^-]},$$

where

$$K_o = \frac{[IS_2O_8^{3-}]\,[I^-]}{\{[S_2O_8^{2-}] - [IS_2O_8^{3-}]\}\,[I^-]}$$

comparing equations (4) and (5)

$$k'_o = \frac{k_{20}\,[I^-]\,K_o}{1 + K_o\,[I^-]}$$

which can be rearranged as

$$1/k'_o = \{1/k_{20}K_o\,[I^-]\} + 1/k_{20}. \tag{6}$$

In presence of micelles and ME, a fraction of $[IS_2O_8^{3-}]$ formed $[F_o]$ will be partitioned towards micelle association $[M\text{–}IS_2O_8^{3-}]\,[F_m]$, resulting in micellar catalysis through concomitant reactions II(a) and II(b). The contributions of $[IS_2O_8^{3-}]$ and $[M\text{–}IS_2O_8^{3-}]$ on k'_m may be written as

$$k'_m = F_o\,k_{20} + F_m\,k_{2m} \tag{7}$$

where fractions

$$F_m = \frac{[M\text{–}IS_2O_8^{3-}]}{[IS_2O_8^{3-}]};$$

$$F_o = \frac{[IS_2O_8^{3-}] - [M\text{–}IS_2O_8^{3-}]}{[IS_2O_8^{3-}]} = 1 - F_m$$

$$k'_m = k_{20} + F_m(k_{2m} - k_{20}).$$

Also

$$K_m = \frac{[M\text{–}IS_2O_8^{3-}]}{[S_2O_8I^{3-}]\,[M]}$$

and

$$K_m = F_m/[M].$$

Rearrangement leads to

$$\frac{k'_m - k_{20}}{k_{2m} - k_{20}} = K_m\,[M]$$

$$= K_m\,[CD - CMC]/N, \tag{8}$$

where CD = surfactant concentration; CMC of SDS = 1.35×10^{-3} mol dm^{-3}; N = SDS micelle aggregation number = 95.0 (ref. 12).

Also

$$\frac{k'_m - F_o k'_o}{F_m} = k'_m,$$

and

$$\frac{1}{k'_m} = \frac{1}{k_{2m}\,K_m\,[I^-]} + \frac{1}{k_{2m}}. \tag{9}$$

The double reciprocal plot equations (6) and (9) ($1/k'_o$ or $1/k'_m$ vs $1/[I^-]$) can be used to determine (k_{20}, K_o) and (k_{2m}, K_m) values.

The extent of micelle/ME catalysis depends on the $[M\text{–}IS_2O_8^{3-}]$, which depends on the available number of free ion-binding sites on the stern layer of the micelle/ME. If the fraction of Na^+ counterion bindings per micelle is $(1 - \alpha)$, then α will be the fraction of

free sites per micelle. Also α equals $[M\text{–}IS_2O_8^{3-}]/[M]$, using K_m and K_o expressions, α can be expressed as $\{K_mK_o[I^-][S_2O_8^{3-}]\}$. At constant $[I^-]$, $[S_2O_8^{2-}]$, α values for normal, and in presence of SDS, ME with various alkanols and oils may be obtained and compared. The catalytic activity coefficient β^{-1} equal {vII(b)/vII(a)} which equals $(k_{2o}/k_{2m}) \cdot \alpha/K_m \cdot K_o$ may be obtained.

Experimental

Materials

SDS, $K_2S_2O_8$, KI, $Na_2S_2O_3$ (Lobachemie) and starch (E Merck) were used as such. Benzene, toluene, xylene (E Merck) were distilled fresh before use.

Kinetic measurements

The kinetics were followed by measuring the clock time of appearance of blue colour (I_2 generation) from the mixture containing KI, $K_2S_2O_8$, $Na_2S_2O_3$, and starch; and SDS or ME if necessary[14]. The reactants were N_2 purged and thermostated at desired temperature. The reaction time commenced from the addition of $K_2S_2O_8$.

Microemulsion preparation

To an aqueous solution of SDS, the oil for 5% of H_2O was added and the turbid solution was titrated with constant stirring against the *n*-alkanol taken in a microburette until transparency was reached. The pseudo-ternary phase diagrams encompassing the oil in water microemulsion region correspond to the percentage compositions of the microemulsion constructed. The phase boundaries are checked by preparing above and below compositions and tested for ME properties. ME medium was N_2 purged and the weighed quantities of the clock reaction components were added and the time of appearance of I_2 was noted.

During the reaction time, ME were found stable and no phase separation resulted. Changes in oil and cosurfactant nature were monitored by adding the respective components and maintaining the similar conditions. During oxidation reactions, alcoholic oxidation did not occur.

Electronic spectra of pure SDS, KI, $K_2S_2O_8$ solutions and mixture of KI and $K_2S_2O_8$, KI + $K_2S_2O_8$ + SDS are recorded in a UV Specord Carl Zeiss Spectrometer in the 200 to 350 nm range.

Results and discussion

Figure 1 shows pseudoternary phase diagram of benzene, toluene and xylene containing SDS–water–*n*-alkanol microemulsions. Xylene produces extended L_1 microemulsion boundaries than benzene and toluene. Higher solubilization of xylene may be accounted for branched cosurfactant like isopropanol. Generally higher solubilization causes interfaces with higher extent of fluidity, lower extent of counterion binding and higher substrate association.

Table 1 gives the phase compositions of four components of oil/water ME employed in the kinetics. The compositions fall within the ME domains in the pseudo-ternary phase diagram of O/W regions (Figure 1). During the study of salt effect, there is no phase separation in these formulations due to the low concentration of the salt and possibly due to the counterion-binding effects of the double layer.

Table 2 gives the overall pseudo-first-order rate constant values (k) at different [substrate] under normal, micellar and ME conditions. Increase in T increases k' values. Separate rate constants for $S_2O_8^{2-}$ decomposition (k_d) and I^- oxidation (k_o'), determined using equation (3) and Figure 2 for the uncatalysed and catalysed reactions, are given in Table 3. The linear fits of double reciprocal plots given in Figure 4 [equation (6) and equation (9)] are used to determine the values of (k_{2o}, K_0) and (k_{2m}, K_m) for the I^- oxidation process in absence and presence

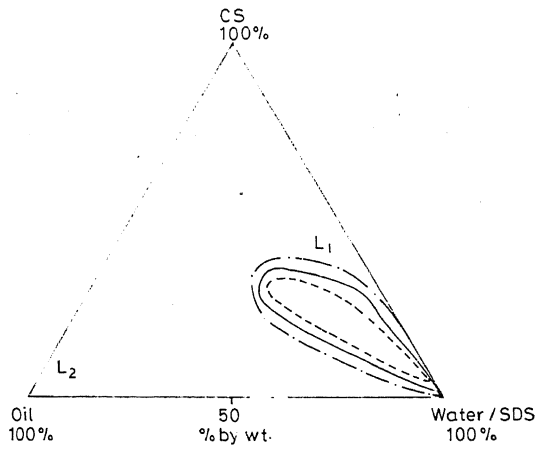


Figure 1. Pseudoternary phase diagram of oil–SDS–water–*n*-alkanol microemulsions, with L_1 regions at 27°C. –·–, xylene–isopropanal ME; —, benzene–propanol ME; – – –, toluene–propanol ME.

Table 1. Percentage composition of the components in oil–water–SDS–*n*-alkanol microemulsions

Oil type	Cosurfactant type	% Wt oil	% Wt water	% Wt SDS	% Wt cosurfactant
Benzene	Propanol	2.252	76.90	0.148	20.66
Toluene	Propanol	2.133	73.89	0.142	23.82
Xylene	Propanol	2.108	71.87	0.138	25.87
Benzene	Isopropanol	2.175	74.30	0.143	23.36
Toluene	Isopropanol	2.005	69.46	0.133	28.39
Xylene	Isopropanol	1.886	64.32	0.123	33.70

Table 2. Effect of [reactant] on the overall pseudo-first-order rate constant of the parallel reactions comprising I^- oxidation and $S_2O_8^{2-}$ decomposition, in presence (k_m'') and absence (k_o'') of SDS micelles and ME

	Temperature (°C)					
	27			60		
$[S_2O_8^{2-}]$ M	$k_o'' \times 10^3$ (s^{-1})	$k_m'' \times 10^3$ (s^{-1})	$k_{ME}'' \times 10^3$ (s^{-1})	$k_o'' \times 10^3$ (s^{-1})	$k_m'' \times 10^2$ (s^{-1})	$k_{ME}'' \times 10^2$ (s^{-1})
1.43	3.09	3.90	5.43	6.69	1.22	6.97
1.82	3.48	4.88	6.26	8.31	1.66	8.15
2.17	3.90	5.80	7.48	9.50	2.02	10.64
2.50	4.02	6.79	8.21	11.41	2.62	13.93

$[I^-] = 10\,[S_2O_8^{2-}]$ M; $[SDS] = 9.09 \times 10^{-3}$ M.
ME = Water–SDS–toluene–*n*-propanol.

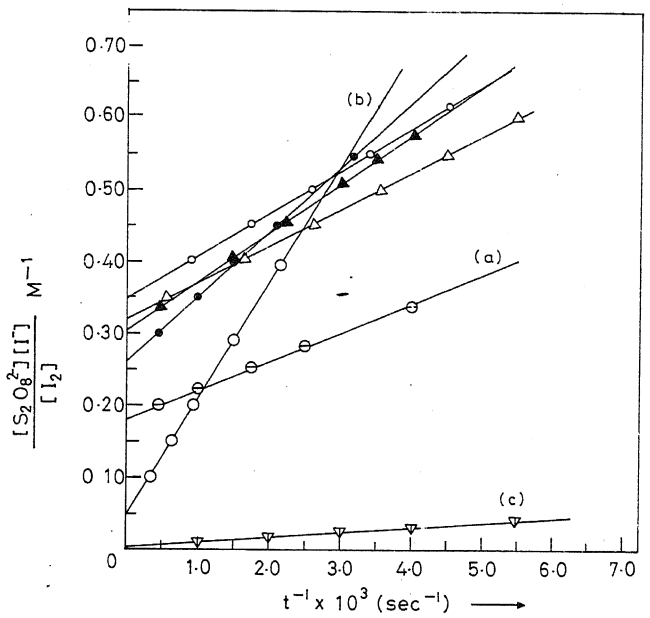


Figure 2. Overall kinetic plots of I^- oxidation by $S_2O_8^{2-}$ in the normal (▲) and presence of SDS micelles (O) and microemulsions and in Na_2SO_4 (Δ) at RT and (●) at 60°C; Me = oil (*a, b, c*)–SDS–water–isopropanol, *a* = benzene; *b* = toluene; *c* = xylene.

of micelles and microemulsions. These values are given in Table 4. The effect of surfactant concentration on k_m is shown in Figure 3. Even though rate constant values are altered, the orders of the parallel reactions (first order decomposition and second order oxidation) remain unaltered in presence of micelles/ME.

Study of the rate constant values shows some salient features: addition of Na_2SO_4 enhances the overall rate constant in accordance with the ionic strength effect on ionic reactions containing the same sign ions[15]. SDS in aqueous solution may be considered to act as 1 : 1 electrolyte below its CMC. As the concentration of SDS increases while approaching its CMC value, grouping of dodecylsulphate anions for the purpose of micellization occurs. This process causes a sudden depletion of the

Table 3. Pseudo-first-order rate constants of decomposition (k_d) and I^--oxidation (k') values in the normal and presence of SDS micelles and ME

	$k_o' \times 10^2$ (s^{-1})		$k_d' \times 10^3$ (s^{-1})	
	Temperature (°C)			
	30	60	30	60
Normal	1.60	1.77	4.87	6.10
SDS	2.01	4.18	6.98	9.07
Na_2SO_4	1.66	9.40	5.39	24.47
Benzene ME*,[a]	3.14	2.22	5.76	–
Toluene ME*,[a]	5.73	1.25	2.58	–
Benzene ME*,[b]	0.22	3.45	–2.03	–0.30
Toluene ME*,[b]	1.09	1.77	0.77	0.241
Xylene ME*,[b]	0.54	1.75	–1.23	0.55

$[I^-] = 10\,[S_2O_8^{2-}]$; $[Na_2SO_4] = [SDS] = 9.09 \times 10^{-3}$ M.
*Values correspond to 40°C.
[a]2-propanol ME; [b]1-propanol ME.

anions and lowering of ionic strength around the CMC value which may bring in a slight decrease in the rate near the CMC. With further increase in surfactant concentration, the number density of micelles increases and micellar catalysis ensues. Therefore the rate starts increasing from CMC (Figure 3). The creation of appropriate interfaces due to micelles or microemulsion enhances the substrate interfacial bindings and decreases the approachability barrier between similarly charged substrates. Here, this effect is more pronounced, since binding cum exchange of sulphate ions of $S_2O_8^{2-}$ in the medium and the sulphate ions of the SDS micelle stern layer is more probable[16]. If it is so, the micelle/ME-interfacially associated substrate/intermediate would probably create a concomitant similar to the normal reaction.

In the phase components belonging to outside boundary and falling within the bicontinuous phase or any other heterogenous phase lowers the rate due to impeding effect of the other phases. When other compositional variation that lead to other phase changes like bicontinuous microemulsions is tried for reaction media, the

Table 4. Association constant and rate constant of I^- oxidation by $S_2O_8^{2-}$ in the normal and presence of SDS micelles and ME $[Na_2SO_4] = [SDS] = 9.09 \times 10^{-3}$ M

	Temperature (°C)			
	30		60	
	$k_{2m} \times 10^2$	K_m	$k_{2m} \times 10^2$	K_m
Normal	0.74	47.76	2.20	15.91
SDS	0.91	21.49	7.93	15.56
Na_2SO_4	0.70	5.00	–	–
Benzene ME*[a]	0.93	27.67	1.50	15.85
Toluene ME*[a]	0.72	10.31	–	–
Xylene ME*[a]	0.43	4.79	–	–
Benzene ME*[b]	0.17	2.83	–	–
Toluene ME*[b]	0.25	2.57	–	–

Normal, k_{2m} and K_m correspond to k_{20} and K_o.
*Values correspond to 40°C.
[a]2-propanol; [b]1-propanol.

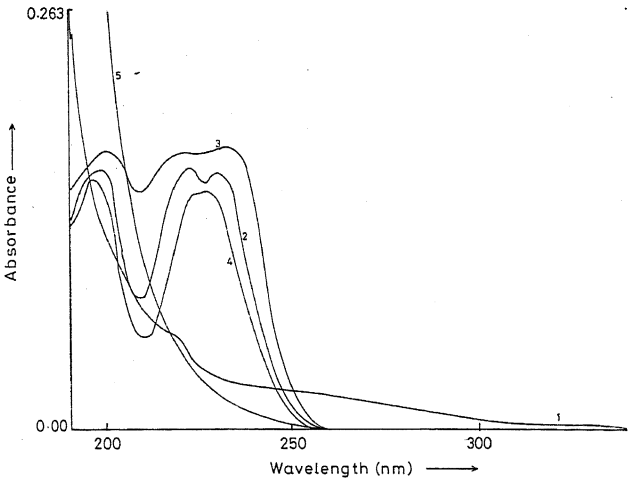


Figure 5. UV spectra of aqueous solution of pure SDS (1); $KI + K_2S_2O_8$ (2); KI (3); $KI + K_2S_2O_8 + SDS$ (4); $K_2S_2O_8$ (5) at 25°C.

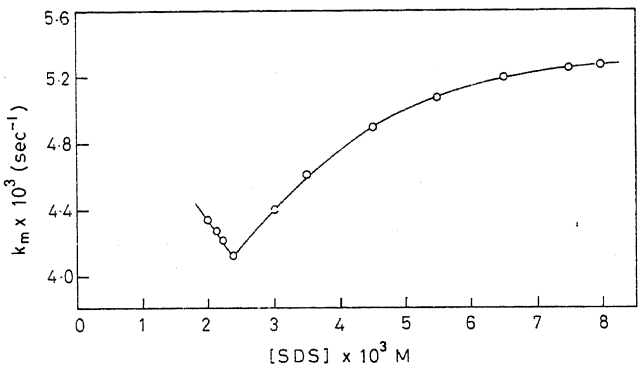


Figure 3. Pseudo-first-order rate constant (k_m) dependance on SDS concentration at 27°C.

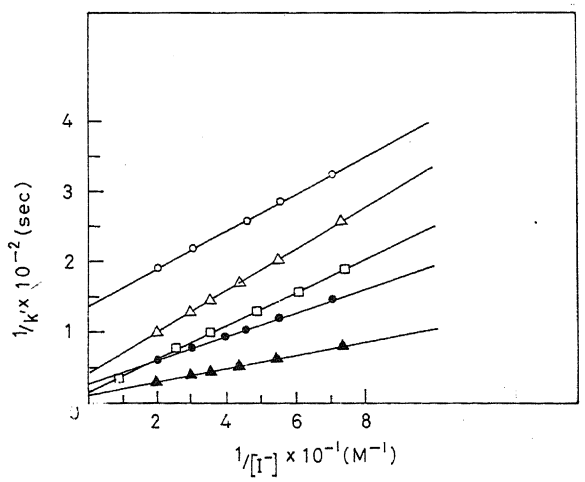


Figure 4. Double reciprocal plots of k_{20}^{-1} (O, □), k_{2m}^{-1} (Δ) vs $[I^{-1}]^{-1}$ for uncatalysed (O), in presence of Na_2SO_4 (□) and micelle catalysed iodide oxidation reactions at 27°C (O, Δ) and 60°C (●, ▲).

Table 5. Fraction of free counterion binding of SDS micelles and ME (α) and catalytic activity coefficient (β), E_a and ΔH values of I^-–$S_2O_8^{2-}$ reactions

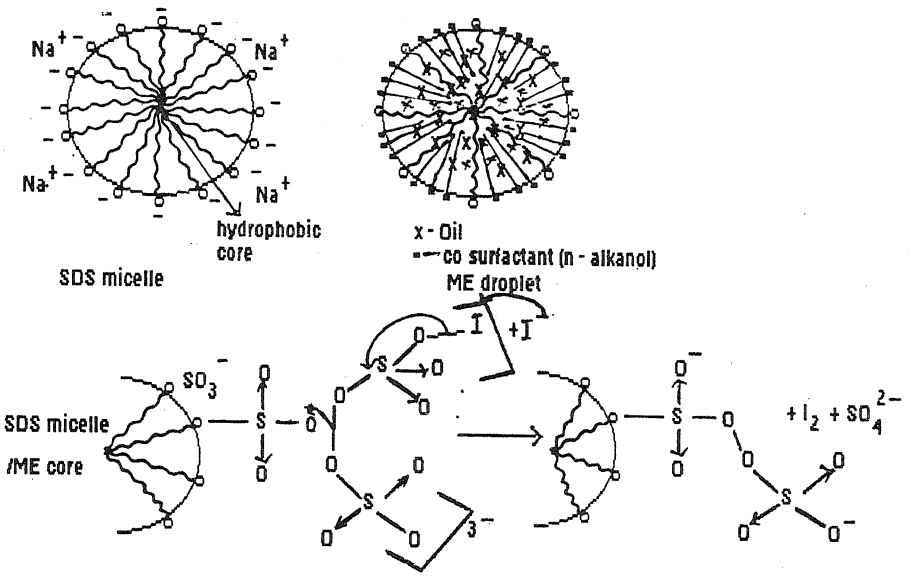
Condition	α	1/β	E_a cals/mole	ΔH cals/mole
Normal	–	–	626.84	131.04 (ΔH_o)
SDS	0.019	0.010	4578.50	379.07 (ΔH_m)
Benzene ME[a]	0.021	0.028	–6509.67	135.46
Toluene ME[a]	0.024	0.033	–28655.50	–
Xylene ME[a]	0.029	0.039	–	–
Benzene ME[b]	0.025	0.034	–	–
Toluene ME[b]	0.022	0.030	–	–

$[I^-] = 10\,[S_2O_8^{2-}]$ M; $[SDS] = 9.09 \times 10^{-3}$ M.
[a]2-propanol; [b]1-propanol.

reactions exhibited low rates and the reactivity difference may be due to the structural variation in surfactant assemblies in solutions. In these compositions, the heterogeneity is more due to the higher compositions of O/W ME. So, one may say that there is a decreased reactivity of the reactants in other regions compared to O/W microemulsions.

The values of k_d and k_o' are observed to be accelerated in presence of SDS and microemulsions. In all instances, I^- oxidation seems faster than $S_2O_8^{2-}$ decomposition, since high energy peroxobond needs to be dissociated in the decomposition step.

The UV spectroscopic evidence was sought to visualize the formation of M–$IS_2O_8^{2-}$. When the spectra of separate solutions of SDS, KI, $K_2S_2O_8$ are compared with that of the mixture containing micelle, I^-, $S_2O_8^{2-}$, then the characteristic λ_{max} of $S_2O_8^{2-}$ seems to have shifted. The shift is significant when compared with pure SDS micelles towards red shift. This can be attributed to the interfacial substrate binding. However, this peak variation



Scheme 1.

with time could not be monitored and approaches as (2). This spectrum (4) was obtained instantaneously after mixing.

The extent and the stability of the substrate association with the micelle/ME interfaces are reflected from association constant values (K_o, K_m). $K_o > K_m$ trend indicates that presence of SDS/ME aid the decomposition of the intermediate by forming a lesser stable complex. Lower K_m values are also due to the lesser probability of binding of anionic intermediate on anionic SDS micelle interface. However concomitant fractionation of $IS_2O_8^{3-}$ to $M–IS_2O_8^{3-}$ leads to an overall acceleration of the rate. The interfacial interactions are shown in scheme 1.

Regarding microemulsions, solubilization of small proportions of oil occurs at hydrophobic micellar core, resulting in oil-swollen micelles. Addition of cosurfactants may sometimes result in further fractionation of droplets. In any case there is an excess interfacial area created during microemulsification. This effect increases the fraction of free ions at the stern layer of the micelle and the oil-swollen micelle. Therefore interfacial association of the substrate increases. Also, due to the enhanced interdroplet interactions, certain amount of instability prevails on the $[M–IS_2O_8^{3-}]$. The presence of cosurfactants tends to lower the extent of micellar catalysis since the cosurfactant molecule adheres competitively at the interface. Due to a considerable reduction in the interfacial effect, $k_{2m} < k_{20}$ is noted. With increase in the carbon chain length of the cosurfactant, k_{2m} values are lowered. The overall k values increase from benzene, toluene and xylene while k_d values suffer a decrease since phase separation of ME occurred occasionally, the data remain nil in such cases.

In ME the fraction of free micellar ionic sites increases, i.e. α-values increase in ME (Table 5). β^{-1} values measure the extent of catalysis and the role of interfacial effect. Even though α increases, β decreases in certain cases due to the resistance offered by the cosurfactant on the $M–IS_2O_8^{3-}$ product conversions.

Increase in temperature increases the k, while K decreases. The thermal effects induce instability of the associated complex due to competitive desorption of the substrate from the interface. But due to the lowering of K values, k_{20} and k_{2m} values increase, i.e. instability of the associated intermediate is removed through conversions leading to products. The E_a and ΔH values of the catalysed and normal reactions show negative values when an increase in temperature decreases the k and K values. This effect is large in case of ME. The ΔH values of the normal and ME media are comparable while that of SDS is more. The close packing of the surfactant molecule in the aggregates is loosened in the ME due to the presence of cosurfactants and oil inclusions. Also, micelle and ME droplet breakdown processes prevail when T increases.

Thus creation of interfaces inside the bulk medium through the presence of micelles or ME certainly induces acceleration of the decomposition and the iodide oxidation processes of $S_2O_8^{2-}$ ions. The separation of the rate constants of the individual step from the overall values applies for the normal and micelle/ME media. However, the collisional and diffusional properties of micelles and ME stand to be probed for future studies.

1. López, M. A.-Quintela and Losada, D., *Phys. Rev. Lett.*, 1988, **61**, 1131–1134.
2. Robinson, B. H., *Chem. Br.*, 1990, **26**, 342.

3. Lianos, P. and Thomas, J. K., *J. Colloid Interface Sci.*, 1987, **117**, 505.
4. Fletcher, P. D. I. and Parrot, J., *J. Chem. Soc. Farady Trans. I*, 1988, **84**, 1131.
5. Moya, L. M., Isquierdo and Casado, *J. Phys. Chem.*, 1991, **95**, 6001.
6. Munoz, E., Gomez-Herrera, C., del Mar Graciani, M. and Moya, L. M., *J. Chem. Soc. Faraday Trans.*, 1991, **87**, 129.
7. Larrson, M. M., Patrick, Aldercruetz and Mattiasson, B., *J. Chem. Soc. Faraday Trans.*, 1991, **87**, 465.
8. Overbeck, J. T. C., De Bruyn, P. L. and Verhoeck, F., in *Surfactants*, Academic Press, London, 1984.
9. Salero, C., Lucano, A. and Fasella, D., *Biochemistry*, 1989, **71**, 461.
10. House, D. A., *Chem. Rev.*, 1962, **62**, 185.
11. Kolthoff, I. M. and Carr, E. M., *Anal. Chem.*, 1953, **25**, 298.
12. Fendler, E. J. and Fendler, J. H., *Catalysis in Micellar and Macromolecular Systems*, Academic Press, New York, 1975, pp. 86–100.
13. Bamford, C. H. and Tipper, C. F. H., *Comprehensive Chemical Kinetics*, Elsevier, Amsterdam, 1972, vol. 6, p. 352.
14. King, C. V. and Jacobs, M. B., *J. Am. Chem. Soc.*, 1931, **53**, 1705.
15. Laidler, (ed.), *Chemical Kinetics*, Harper and Row, London, 1987, 3rd edn, Chap. 6, pp. 197–200.

ACKNOWLEDGEMENT. We thank UGC, New Delhi, for financial assistance.

Received 5 April 1995; revised accepted 22 September 1995

RESEARCH COMMUNICATIONS

Some FRW models with constant active gravitational mass

Abdussattar and R. G. Vishwakarma

Department of Mathematics, Banaras Hindu University, Varanasi 221 005, India

The consequences of taking the active gravitational mass of the universe constant in the background of FRW models are investigated. It is found that the dependence of the nature of expansion on the curvature parameter k may be altered.

IN the standard big bang cosmology, wherein the universe is assumed to be filled with a distribution of matter represented by the energy momentum tensor of a perfect fluid

$$T_{ij} = (\rho + p)v_i v_j + p g_{ij}, \quad \text{(in the units with } c = 1) \qquad (1)$$

and the geometry of the universe is described by the Robertson–Walker line element

$$ds^2 = -dt^2 + R^2(t)\{(dr^2/1 - kr^2) + r^2(d\theta^2 + \sin^2\theta\, d\phi^2)\}, \quad k = \pm 1, 0, \qquad (2)$$

the Einstein field equations

$$R_{ij} - \tfrac{1}{2} R_h^h\, g_{ij} = -8\pi G T_{ij} \qquad (3)$$

obtain two independent equations

$$-\ddot{R}/R = (4\pi G/3)(\rho + 3p), \qquad (4)$$

and

$$\dot{R}^2/R^2 + k/R^2 = (8\pi G/3)\rho. \qquad (5)$$

The expression for the scale factor R is obtained by solving the differential equation resulting from equations (4) and (5) by assuming an equation of state. Out of the three cases of the model obtained for different values of the curvature parameter k, one is closed for $k = +1$ and the remaining two are ever expanding for $k = 0$ and -1. In this communication, motivation is given for taking the active gravitational mass of the universe as constant and it is found that the dependence of the nature of expansion of the resulting models on the curvature parameter k may be different from that of the standard models.

The field equation (4) may be deemed as an analogue of the Newtonian force law and suggests that the force per unit mass at each space-time point is determined by the active gravitational mass density $(\rho + 3p)$. Here we note that the gravitational pull is exerted not only by ρ as in the Newtonian theory but rather by $(\rho + 3p)$ which exhibits a relativistic effect. It is this additional pressure and internal energy contribution to the gravitational force which is the major cause of the problem of gravitational collapse in general relativity[1].

One notes that equations (4) and (5) obtain $\rho R^3 = $ constant (ref. 2) in the present pressureless phase of evolution, which can be interpreted as the conservation of the total active gravitational mass of the comoving sphere of radius R. As there is no justification of conferring a special status upon the present epoch, we speculate that this constancy feature of the active gravitational mass is met not only in the present phase of evolution but in the early phases too. We thus assume that in the presence of pressure, the active gravitational mass is constant.

$$(\rho + 3p)R^3 = \text{constant} = A \quad \text{(say)}. \qquad (6)$$

Equation (6), taken together with the energy momentum conservation equation

$$\dot{\rho} + 3(\rho + p)\, \dot{R}/R = 0\,, \tag{7}$$

which may be obtained from equations (4) and (5), leads to

$$p^3 = K(\rho + 3p)^2, \quad K = \text{constant} \geq 0\,, \tag{8}$$

which is the equation of state for the perfect fluid constituting the matter content of the universe. It is a physically reasonable equation of state since

$$(dp/d\rho) = [2/3 \cdot (p/(\rho + p))]\,, \tag{9}$$

indicating that $(dp/d\rho) < 1/3$ for $\rho > p$.

Equations (4) and (6) obtain

$$\ddot{R} = -(4\pi GA/3R^2)\,, \tag{10}$$

which, by integrating, gives

$$\dot{R}^2 = (8\pi GA/3R) + B\,, \tag{11}$$

where B is some constant of integration. Equations (5), (6) and (11) may be used to obtain

$$\rho = A/R^3 + 3(B + k)/8\pi GR^2 \tag{12}$$

and

$$p = -(B + k)/8\pi GR^2\,, \tag{13}$$

which are in agreement with equation (8) and suggest that $K = -(B + k)^3/(8\pi G)^3 A^2$ for the self-consistency of the system.

Equation (13) indicates that in order to have $p \geq 0$, one must have $(B + k) \leq 0$.

We note from equation (6) that for $\rho > 0$ and $p > 0$, $A > 0$. Equation (10) hence indicates that $\ddot{R} < 0$. This, together with $\dot{R} > 0$, implies that the curve $R(t)$ must have reached $R(t) = 0$ at some finite time in the past, say at $t = 0$, which is a singularity, where equation (6) holds in the limiting case. The models thus start from a big bang.

Equation (11) may be integrated for different values of k as:

$k = 0$

We thus have $B \leq 0$ for $p \geq 0$. When $B < 0$ say $B = -n^2$, we obtain

$$t = -\frac{(24\pi GAR - 9n^2R^2)^{1/2}}{3n^2} + \frac{4\pi GA}{3n^3} \times \sin^{-1}\left[\frac{3n^2R}{4\pi GA} - 1\right] - \frac{2\pi^2 GA}{n^3}\,. \tag{14}$$

When $B = 0$, we have

$$R = (6\pi GA)^{1/3}\, t^{2/3}\,. \tag{15}$$

Equation (14) indicates that the model reaches its maximum radius $R_{max} = 8\pi GA/3n^2$ at $t = 4\pi^2 GA/3n^3$ and thereafter starts contracting back to origin in a way similar to the $k = 1$ case of the standard model though here $k = 0$. Equation (15) suggests that $R \to \infty$ as $t \to \infty$ with the deceleration parameter $q = \frac{1}{2}$ always as in the $k = 0$ case of the standard model.

$k = -1$

In this case $p \geq 0$ demands $B \leq 1$. If $B > 0$, say $B = n^2 \leq 1$, we obtain

$$t = \frac{(24\pi GAR + 9n^2R^2)^{1/2}}{3n^2} - \frac{4\pi GA}{3n^3} \times \cosh^{-1}\left[\frac{3n^2R}{4\pi GA} + 1\right]\,, \tag{16}$$

which indicates that $R \to \infty$ and $\dot{R} \to n$ as $t \to \infty$. The deceleration parameter, which is initially $\frac{1}{2}$ in this case tends to zero as $t \to \infty$. For $B \leq 0$, the evolution is described by equations (14) or (15).

$k = 1$

In this case, $p \geq 0$ demands $B \leq -1$ and the evolution is governed by equation (14) which describes a closed universe as mentioned earlier.

We thus see that the dependence of the nature of expansion on the curvature parameter k may be altered by the incorporation of the conservation law (6) in the standard cosmology.

1. Ellis, G. F. R., in *General Relativity and Cosmology* (ed. Sacks, R. K.), Academic Press, 1971, pp. 127.
2. Weinberg, S., *Gravitation and Cosmology*, Wiley, New York, 1972, pp. 472.

Received 25 August 1995; accepted 17 October 1995

Evaluation of intra-specific variability in *Avicennia marina* Forsk. using RAPD markers

P. Balakrishna

M.S. Swaminathan Research Foundation, Madras 600 113, India

Random amplified polymorphic DNA (RAPD) markers for determining the genetic diversity among Mangrove species, *Avicennia marina* Forsk. were used. Individuals collected from several sites in India were analysed for their genetic relationship. Both inter- and intra-populational level polymorphisms were observed. The dendrogram arising out of the RAPD data shows prominent patterns of relationships that are discussed in the light of the respective habitats of the species distributed in India. The usefulness of RAPDs in biodiversity assessment and evaluation has therefore been tested.

UNDERSTANDING the genetic basis of diversity of plants helps utilize plant genetic resources effectively. Because of the long-term life cycle, tree genetics tends to receive less attention than annual crops. Among tree species, mangrove forest trees have not so far been subjected to detailed genetic analysis. Mangroves form an important group of plants for several of their characteristics, including their ability to withstand high levels of salinity, besides protecting coastal areas from seawater intrusion[1]. However, this group has suffered from lack of information on their genetics. They are distributed all over tropics and 59 species belonging to 29 families occur in India. Mangrove habitats now face serious threat due to extensive felling, changing coastal environment including lack of freshwater inflow[2]. The performance of each species differs in such an environment and differences are apparent[3]. Conservation of mangroves is thus a major priority in coastal area development. Of several cosmopolitan species of mangroves, *Avicennia marina* is an important species because it can withstand varying soil conditions, high levels of salinity, lack of freshwater inflow for considerable periods of time and pollution of varying forms[4]. Besides they have good rates of regeneration and show different adaptational features. Genetic characterization of mangroves is unfortunately limited owing to lack of concerted studies on identifying the genetic variability within and in-between species. There is no report on the genetic diversity of mangroves except for *Rhizophora* species[5]. Though isozyme studies have been carried out in mangroves their reproducibility is rather poor.

Genetic analysis of tree species, especially of forest trees, has not received much attention from molecular biologists until recently. Relatively few genetic maps are available for forest trees[6–8]. This is mainly due to lack of suitable pedigrees and basic genetic data like number of chromosomes, linkage groups and even population genetic analysis.

Mangroves are threatened all over the world today due to direct and indirect exploitation. Conservation of mangroves has thus become important not only to protect the coastal areas but also to protect the coastal communities from seawater intrusion and from the adverse impacts of potential changes in sea level rise. One of the suggestions for mangrove restoration is to select genetically superior germplasm and to introduce them into target sites[9]. But for this to happen it is important to understand the extent of diversity in individual species and the relationship these species have with their respective habitats. Diversities in structural formations of mangrove ecosystems can be witnessed along the latitudinal and longitudinal gradients. This may be due to the climate or soil or due to water availability[10]. Thus there is a necessity to evaluate the genetic diversity of these trees.

Random amplified polymorphic DNA (RAPD)[11] markers have been successfully used to assess the genetic diversity within several groups of plants, including tree species like conifers, Norway spruce, pine and others[12–14]. Besides this, RAPDs were also used for several other purposes like genome mapping, population genetics, introgression analysis[15]. RAPDs are better choice of markers for evaluation of genetic diversity in species where cloned markers are not available. Hence RAPD analysis was undertaken in mangrove species, *A. marina* which showed clear inter- and intra-populational variability from the collections made in India. The sites selected were from different habitats ranging from unpolluted Pichavaram to polluted sites like Ennore with petrochemical pollution and Adyar with inorganic pollution. Duplicate samples were collected from Ennore and Bombay so that they can represent the trueness of the genetic analysis. Genetic correlations and genetic distance analysis were done using RAPD data to relate them with the habitat and growth.

A. marina accessions were collected from different sites in India (from Bombay, Ganapatipule in Maharashtra, Pichavaram in Tamil Nadu, Ennore and Adyar in Madras). Seven accessions representing five locations were chosen to see if the RAPD markers were useful to study their genetic diversity. For detecting the variations a total of 25 primers were initially used, out of which 12 gave diagnostic polymorphisms of inter- and intra-populational variability.

Total genomic DNA was isolated from lyophilized leaf material. To 300–400 mg of powdered leaf tissue, 1000 µl of extraction buffer (1 M Tris-HCl, pH 8.0, 5 M NaCl, 0.5 M EDTA, 14 β mercaptoethanol, and 1% CTAB) was added and incubated at 65°C for 90 min with occasional mixing. The mixture was then cooled

to room temperature and 500 µl of chloroform : octonol (24 : 1) was added. After centrifuging at 10,000 rpm at 4°C for 10 min the supernatant was transferred to a fresh tube. To this 750 µl of ice-cold isopropanol was added. The DNA was washed with 500 µl of 70% ethanol and 750 µl of 3 M sodium acetate and again with 70% ethanol. The DNA was then air-dried and suspended in 1X TE buffer.

Decamer oligonucleotide primers were used in this study. Table 1 lists primers that were selected for the final analysis. Amplifications were done in volumes of 25 µl containing 10 mM Tris-HCl, pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 0.1 mM each of dATP, dTTP, dGTP, and dCTP (Perkin Elmer, USA), 5 ng of genomic DNA, 0.5 µM of primer and 3 units of *Taq* DNA polymerase (Perkin Elmer, USA). The amplification was carried out on a Perkin Elmer 480 thermal cycler (Perkin Elmer, USA). Samples were amplified at one step of 2 min at 94°C, 40 repeats of the following cycles—one min at 94°C, one min at 37°C, two min at 72°C, and then a final extension time of 4 min at 72°C. The amplified products were then analysed on 1.8% agarose gels stained with ethidium bromide.

For each primer, bands were scored for their presence (+) or absence (–) of a amplified product. From these data a simple matching coefficient (SMC) analysis was carried out on a GENSTAT 5 program (licensed to Lawes Agricultural Institute, UK) to analyse the genetic distance and genetic relatedness. SMC takes into account the entire amplification products from various genotypes under study against all the primers scored for. Thus weightage is given as a shared presence or absence of bands also in calculating similarities and distance of different genotypes.

Table 1. List of amplification products generated by using arbitrary 10 base oligonucleotides, based on all accessions

Primer	Sequence	No. of fragments detected: within *A. marina*	No. of fragments detected: between *A. marina*
1.	3′ GTTTCGCTCC 5′	35	19
2.	3′ TGATCCCTGG 5′	14	14
3.	3′ CATCCCCCTG 5′	27	16
4.	3′ GGACTGGAGT 5′	22	13
5.	3′ TGCGCCCTTC 5′	21	12
6.	3′ CCATTCCGAG 5′	2	2
7.	3′ CACCTGCTGA 5′	17	4
8.	3′ GTGAATGCGC 5′	4	1
9.	3′ CCTACACGGT 5′	15	5
10.	3′ CAGCCGAGAA 5′	11	3
11.	3′ GAAGGAGGCA 5′	21	8
12.	3′ TGAACCGAGG 5′	16	6

Isolation of DNA from Mangrove species has been a difficult problem until now. This is because they contain a number of secondary metabolites like phenolics, terpenoids, alkoloids and mucilage. The DNA extraction protocol followed here solved this problem. Of the 25 primers screened, 12 were found to be suitable for diagnosis. Amplified products could be detected from all primers used. A total of 205 amplified products were scored. Within an individual plant the primers detected RAPD patterns containing DNA bands within 0.1 to 2.3 kb.

Except for primer 4 (Table 1) all the others gave good variability profiles. Primer 4 was not able to detect

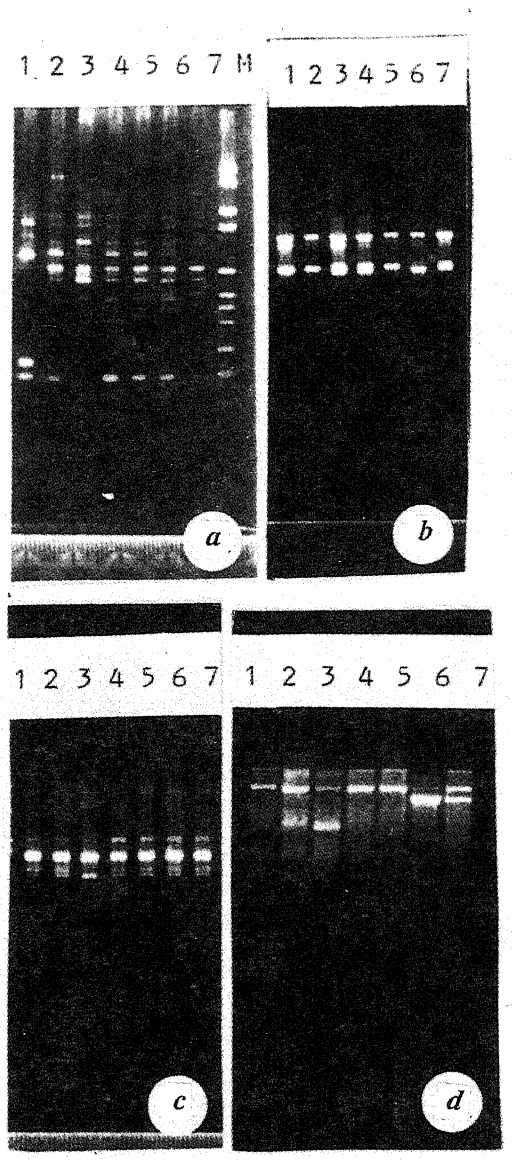


Figure 1. Example of amplified products from *A. marina* populations collected from: Lanes 1 & 2, Bombay; 3, Ganapatipule; 4, Pichavaram; 5, Adyar; 6 & 7, Ennore. ***a***, Amplified products using primer 3′ GTTTCGCTCG 5′; ***b***, Amplified products using primer 3′ TGATCCCTGG 5′; ***c***, Amplified products using primer 3′ CATCCCCCTG 5′; ***d***, Amplified products using primer 3′ GGACTCCAGT 5′.

any variation within the samples. This polymorphism was apparent with all the primers and was not only between samples but also within the samples collected from different sites (Figure 1). The dendrogram constructed based on the SMC analysis and represented in Figure 2 distributes the genotypes into different clusters. The amount of variability detected when the DNA sample was subjected to repeated RAPD analysis was the same, confirming the reproducibility of the analysis.

The two sites selected each in Ennore and Bombay showed considerable variations (Figure 1). All the primers except primers 4 and 8 detected variations within these sites and the dendrograms showed their relatedness (Figure 2). Genetic distance map was also constructed based on the RAPD data using SMC (Figure 3) to identify the distances. It was found that all the genotypes were distinctively different and separate based on amplified products.

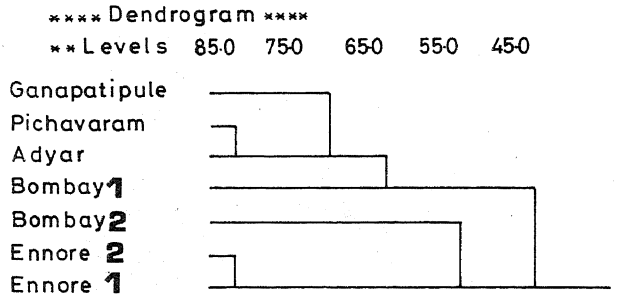


Figure 2. Dendrogram showing genetic relationships between *A. marina* populations constructed using simple matching coefficient analysis based on RAPD analysis.

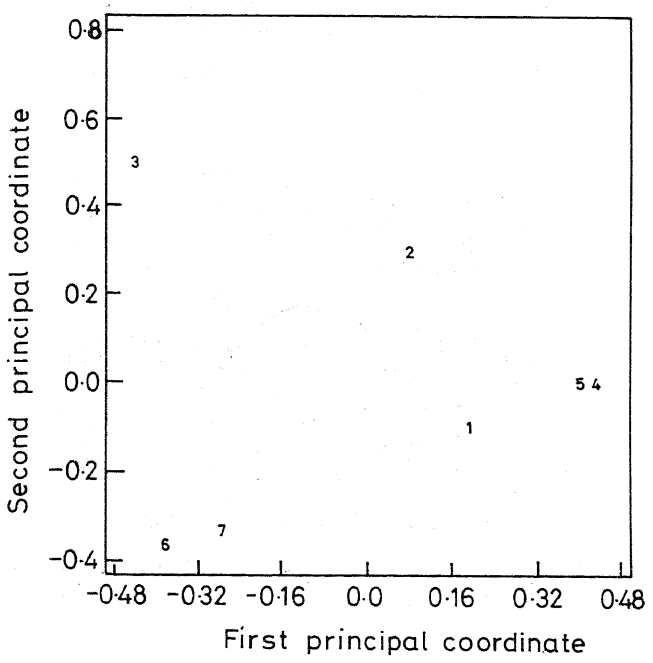


Figure 3. Genetic distance analysis of *A. marina* revealed by RAPD markers and analysed using simple matching coefficient analysis. 1 & 2, Populations from Bombay; 3, Ganapatipule; 4, Pichavaram; 5, Adyar; 6 & 7, Ennore.

With an aim to understand the genetic diversity and the relatedness of mangrove species, *A. marina* was taken as a model. This is because of its several and varied adaptational and distributional patterns. RAPD was analysed on seven populations and the variability patterns were scored on the basis of amplified products. The analysis was carried out not only on a pairwise combination of variability but also accounted for the entire information given by all the genotypes against all the primers used. This comparison is based on the presence or absence of a band (amplified product) and give weightage as a shared presence in calculating similarities[16].

The distribution of the genotypes observed on the dendrogram is represented in Figure 2 and shows that the RAPD data can be a powerful tool in assessing the genetic diversity. With regard to intra-site variability, there is variability within *A. marina* collected from the same sites. This is surprising because the species have evolved from the same habitats for a long period of time. The frequent inundation of seawater might have helped cross pollination to account for this polymorphism. Though the dendrogram is quite revealing the groups are considered different with respect to sites. The similarity between the material collected from Bombay and Ennore may be due to similarities in soil conditions like hypersalinity and the type of pollution[17]. Ganapatipule and Pichavaram are considered to be reasonably less polluted and also received copious amounts of freshwater inflows, while Adyar falls between these two groups. This relatedness is also established in the genetic distance and similarity analyses.

The RAPD technique contributes to a better understanding of genetic characterization than the analysis of isozymes. For developing conservation methodologies it is important to evaluate the genetic diversity within a given species. Intra-populational level variation helps to identify suitable genetic material also. This kind of data *vis-à-vis* the other physiological, edaphic factors are taken into account when identifying suitable clones for introduction into different areas. Also, this study proves the utility of RAPDs in biodiversity measurement.

1. IUCN, in *Commission in Ecological Papers* (eds Hegrel, E. J. and Davie, J. D. S.), Paper 3, 1983, pp. 1–88.
2. Walsh, G. E., in *Ecology of Halophytes* (eds Reinold, J. and Queen, W. H.), 1974, pp. 51–174.
3. Odum, W. E. and Johannes, R. E., in *Tropical Marine Pollution* (eds Wood, E. J. F. and Johannes, R. E.), 1975, pp. 52–62.
4. Scholander, P. F., Hammel, H. T., Hemmingsten, E. A. and Garey, W., *Plant Physiol.*, 1962, **37**, 722–729.

5. Subramonian, N., Ph D thesis, Univ. of Madras, Madras 1993.
6. Neale, D. B. and Sederoff, R., *Can. J. For. Res.*, 1991, **21**, 545–554.
7. Tulsieram, K., Glaubitz, J. C., Kiss, G. and Carlson, J. E., *Biotechnology*, 1992, **10**, 686–690.
8. Gerber, S., Rodolphe, F., Bahrman, N. and Baradat, P. H., *Theor. Appl. Genet.*, 1993, **85**, 521–528.
9. UNESCO, in *The Mangrove Ecosystem: Research Methods* (eds Samuel, C. S. and Snedakar, J. C.), 1984, p. 251.
10. FAO, FAO Forestry paper 117, 1994, p. 319.
11. Williams, J. G. K., Kubelik, A. R., Livak, K. J., Rafalski, J. A. and Tingey, S., *Nucleic Acids Res.*, 1990, **18**, 6531–6535.
12. Carlson, J. E., Tulsieram, L. K., Glaubitz, J. C., Luk, V. W. K., Kauffeldt, C. and Rutledge, R., *Theor. Appl. Genet.*, 1991, **83**, 194–200.
13. Nelson, C. D., Nance, W. L. and Doudrick, R. L., *Theor. Appl. Genet.*, 1993, **87**, 145–151.
14. Binelli, G. and Bucci, G., *Theor. Appl. Genet.*, 1994, **88**, 283–288.
15. Hadrys, H., Balick, M. and Scierwater, B., *Mol. Ecol.*, 1992, **1**, 55–63.
16. Marney, P., Becking, J. R., Hamon, S. and Charriese, A., *Euphytica*, 1994, **74**, 203–209.
17. Selvam, V., Mohan, R., Ramasubramanian, R. and Azariah, J., *Indian J. Mar. Sci.*, 1991, **20**, 67–70.

ACKNOWLEDGEMENTS. We thank Prof. Luigi Monti and his group at the University of Naples, Italy for facilities, to Drs Rob Finch and Robert McNeil, SAC, Scotland for help in computer analysis and to Dr Selvam of MSSRF for his comments on the manuscript. This study was a part of collaborative project supported by Italian Government of MSSRF.

Received 9 December 1994; revised accepted 17 October 1995

Carbon and oxygen isotope trends in late Precambrian–Cambrian carbonates from the Lesser Himalaya, India

B. Kumar and V. C. Tewari*

National Geophysical Research Institute, Hyderabad 500 007, India
*Wadia Institute of Himalayan Geology, Dehradun 248 001, India

$\delta^{13}C$ and $\delta^{18}O$ records of the Late Precambrian–Cambrian (Pc–C) carbonates from the Lesser Himalaya are reported here. The data depict two distinct cycles of ^{13}C maxima–minima and one distinct ^{18}O maxima for these carbonates. We suggest that isotopic variations across Pc–C stages relate to marked changes in the carbon and oxygen fluxes.

CARBON isotope signatures of the Precambrian–Cambrian boundary carbonates have been studied from a number of localities of the world[1–6], but their implications are still controversial. The preservation of original $\delta^{18}O$ records[3] in Pc–C boundary carbonates has also been questioned. We present here the result of a carbon and oxygen isotope study of two marine carbonate successions of Late Precambrian–Cambrian age from the Lesser Himalaya. The isotopic data are evolved in relation to sedimentary carbon budget and atmospheric oxygen level (?) changes etc.

In the northwestern Himalaya, two major carbonate-bearing successions, namely, the Deoban Formation of Lower to Middle Riphean age and the Krol Formation of Late Vendian (Ediacaran) age are present[7–9] (Figure 1). The Deoban Formation contains an approximately 1000 m thick succession of carbonate rocks (stromatolitic dolomites, dolomitic limestones, cherty limestones an oolitic limestones) with intercalated beds of shales. It is followed by a shallow marine sequence of the argillo-siliciclastic deposits known as the Simla/Jaunsar Group which comprises of Mandhali, Chandpur and Nagthat formations. The Simla/Jaunsar Group is overlain by the Blaini Formation of early Vendian (Varanginian) age, consisting of diamictite beds and minor deposits of microbial dolomites, siltstones and shales. The Krol Formation of Vendian age includes an ~2000 m thick succession of stromatolitic dolomites, cherty limestones, shales, sedimentary breccia, oolites and grainstones. It

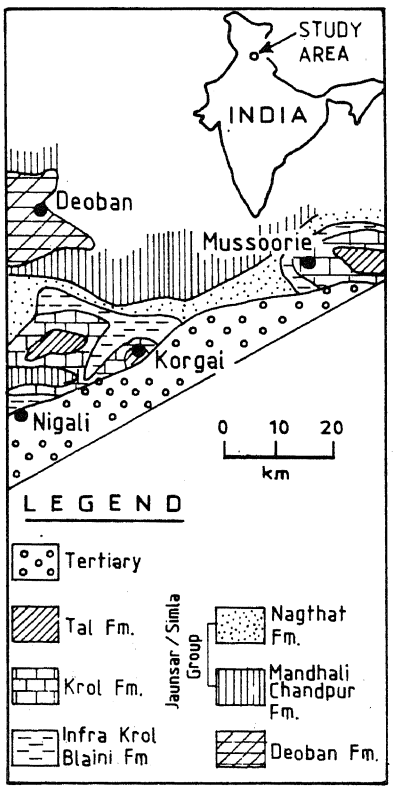


Figure 1. Geological sketch of the Deoban, Blaini and Krol–Tal sedimentary succession of the Lesser Himalaya, showing location of the study area. (map after Chaudhri and Kaltia[22]).

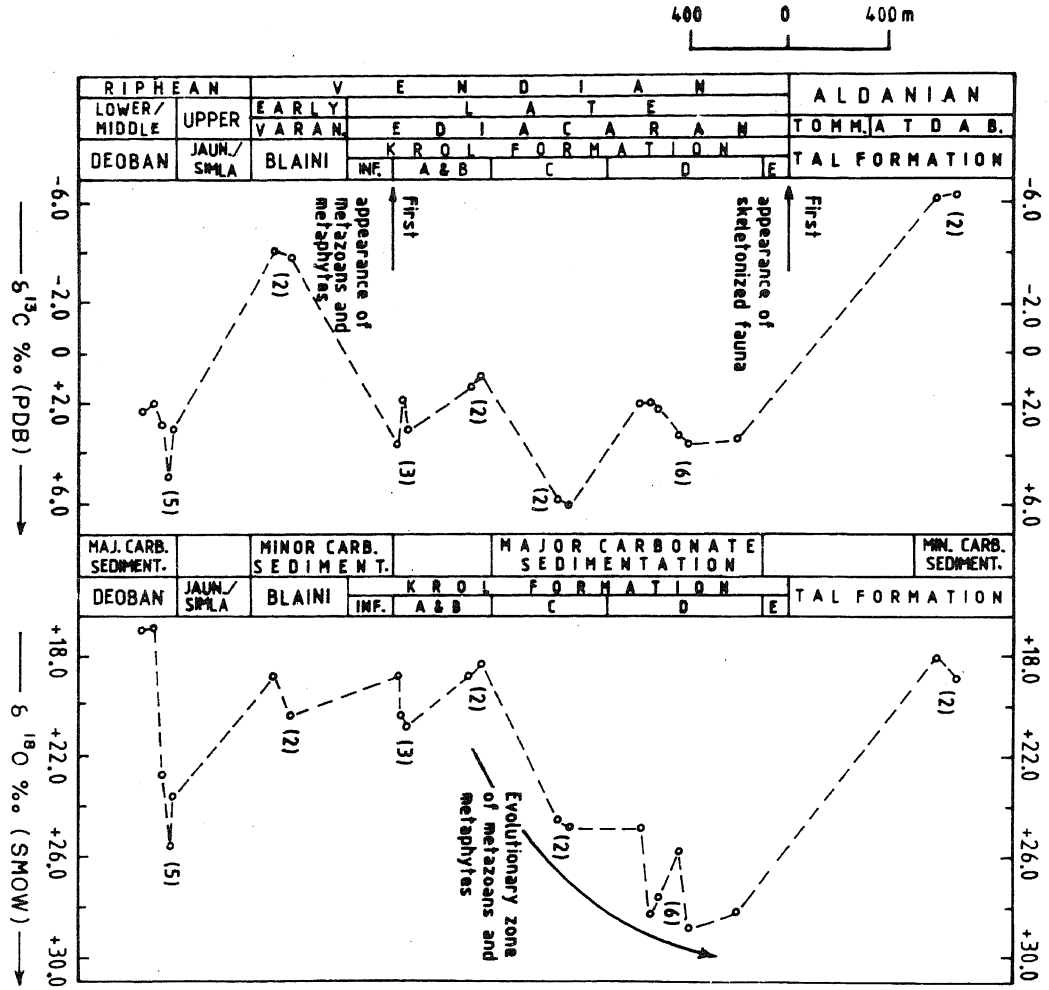


Figure 2. $\delta^{13}C$ and $\delta^{18}O$ records of the Deoban, Blaini, Krol and Tal carbonates in relation to their stratigraphy (the numbers in the bracket refer to number of samples analysed from each formation).

is followed by the Tal Formation (Early Cambrian) which consists of a Chert Phosphorite Member of Tommotian age and argillo–calcareous and siliciclastic member of Atdabanian–Botomian age. The carbonate sedimentation in the Tal Formation is minor. A late Precambrian to early Cambrian age for these formations has recently been established by biostratigraphic studies based on: i) discovery of microfossils of Late Precambrian age[7]; ii) identification of stromatolites of Late Precambrian affinity[8,9]; iii) documentation of Ediacaran metazoans and metaphytes[10,11] in the Lower Krol Member; and iv) records of the early Cambrian small shelly fauna[12,13], trilobites[14], brachiopods[15] and trace fossils[16].

The carbonates investigated in this study have been sampled from the Deoban Formation developed in the Deoban Mountains (30°45′ 77°54′) and those from the Blaini, Krol and Tal formations in the Korgai and Nigalidhar synclines (30°34′50″, 77°39′15″, cf. Figure 1).

$\delta^{13}C$ and $\delta^{18}O$ trends of the Deoban, Krol and Tal carbonates are plotted in Figure 2 in relation to stratigraphy of these carbonates. The isotope trends are drawn through the 'δ' values. $\delta^{13}C$ values become generally more positive from the Middle Deoban reaching ^{13}C maximum of +4.9‰ (PDB) in the Upper Deoban. This is followed by a decline of ^{13}C to a minimum of –4‰ in the Blaini. The carbon isotope profile shows a second rise in the Krol-C depicting a ^{13}C maximum of +6‰ which drops to a minimum of −6.2‰ within the Tal. $\delta^{18}O$ varies from +17.2 to +29.2‰ (SMOW) and exhibits a marked positive shift in the Upper Krol.

The previous workers[5] have opined that the carbon

isotope variations of Krol–Tal sedimentary carbonates of the Lesser Himalaya represent pristine isotopic signatures as these successions have not been subjected to any metamorphism, deformation and excessive burial. We therefore consider that the $\delta^{13}C$ data reported here represent unaltered signatures.

The isotopic trends presented in Figure 2 show that ^{13}C maxima of +4.9‰ and +6‰ (PDB) relate to the major carbonate successions of the Deoban and Krol, and that the ^{13}C minima of −4‰ and −6.2‰ (PDB) are associated with terrigenous clastic deposits with minor carbonates of the Blaini and Tal formations, respectively. The marked increase in carbonate sedimentation during Deoban and Krol times would imply increased fixation of carbon dioxide in the form of carbonate carbon (C_{carb}). Deoban and Krol carbonate sedimentation represents a time span of Riphean to Vendian[7–17], a time interval much larger than the residence time of carbon (~ 10^5 years)[18] in the exogenic cycle. The increased fixation of carbon dioxide in the form of C_{carb} will therefore mean increased availability of CO_2 in the environment existing during Deoban and Krol times. The Deoban and Krol carbonate formations are mainly microbial in nature and their large-scale deposition indicates a relative enhanced build-up of microbial communities which preferentially fixed ^{12}C in the form of organic carbon (C_{org}), this resulting in ^{13}C enrichment in carbonate carbon. ^{13}C maxima thus relate to overall increase in sedimentary carbon ($C_{carb} + C_{org}$) budget, increased availability of carbon dioxide in the prevailing environment, enhanced rate of photosynthesis and possibly warmer climates. Conversely, ^{13}C minima associated with the Blaini and Tal formations may go along with decrease in the total sedimentary carbon budget (as evident from field observations and marked negative $\delta^{13}C$ signatures), reduced rate of photosynthesis, lower concentration of CO_2 in the prevailing environment and possibly colder climates. The colder climates during Blaini times is envisaged by its possible association with the Varangian glaciation event[4]. There is no evidence of colder climates during Tal times.

The $\delta^{18}O$ records of Late Precambrian–Cambrian carbonates are considered to be artifacts of diagenetic alterations[3] or post-depositional isotope exchange equilibration processes[19], but we suggest that this may not be the only possible reason. The Krol carbonates depict a distinct ^{18}O maxima that associate with carbon isotope maxima (Figure 2). ^{13}C maxima has been interpreted in terms of increased carbon burial and enhanced rate of photosynthesis which has been the main source of oxygen build-up in the atmosphere. Further, the Krol sedimentary succession exhibiting distinct ^{13}C and ^{18}O maxima may be related to the evolutionary transition from unicellular life forms to multicellular metazoans and metaphytes, these must have originated in high oxygenated environment[4,20,21]. Therefore, we associate ^{13}C and ^{18}O maxima with intervals of high environmental oxygen level similar to or higher than the present one. However, the causative reason for link between ^{18}O maxima and higher oxygen levels still remains questionable.

It is concluded that $\delta^{13}C$ and $\delta^{18}O$ variations across Pc–C stages relate to the marked changes in the carbon and oxygen fluxes.

1. Hsu, K. J., Oberhansli, H., Gao, J. Y., Shu, S., Haihong, C. and Krahenbuhl, U., *Nature*, 1985, **316**, 809–811.
2. Tucker, M. E., *Nature*, 1986, **319**, 48–50.
3. Magaritz, M., Holser, W. T. and Kirschvink, J. L., *Nature*, 1986, **320**, 258–259.
4. Knoll, A. H., Hoyes, J. M., Kaufmann, A. J., Swett, K. and Lambert, I. B., *Nature*, 1986, **321**, 832–838.
5. Aharon, P., Schidlowski, M. and Singh, I. B., *Nature*, 1986, **327**, 699–702.
6. Magaritz, M., Kirschvink, J. L., Latham, A. J., Zhuravlev, A. Yu. and Rozanov, A. Yu., *Geology*, 1991, **19**, 847–850.
7. Shukla, M., Tewari, V. C. and Yadav, V. K., *Palaeobotanist*, 1986, **35**, 347–356.
8. Tewari, V. C., Proceedings of the 5th Indian Geophytology Conference, Lucknow, 1984, pp. 71–97.
9. Tewari, V. C., *Him. Geol.*, 1989, **13**, 100–135.
10. Tewari, V. C., Abstract in National Seminar on Precambrian Geology, Madras University, 1990, pp. 79–80.
11. Mathur, V. K. and Shankar, R., *J. Geol. Soc. India*, 1989, **34**, 245–254.
12. Ajmi, R. J., *Him. Geol.*, 1983, **11**, 373–409.
13. Bhatt, D. K., Mamagain, V. D., Misra, R. A. and Srivastava, J. P., *Geophytology*, 1983, **13**, 116–123.
14. Kumar, G., Joshi, A. and Mathur, V. K., *Curr. Sci.*, 1987, **56**, 659–663.
15. Tripathi, C., Jangpangi, B. S., Bhatt, B. S., Kumar, G. and Raina, B. K., *Geophytology*, 1984, **14**, 221–227.
16. Singh, I. B. and Rai, V., *J. Palaeontol. Soc. India*, 1983, **28**, 69–80.
17. Valdiya, K. S., in *Precambrian Continental Crust and its Economic Resources* (ed. Naqvi, S. M.), Elsevier, Amsterdam, 1990, pp. 525–553.
18. Holland, H. D., *The Chemistry of the Atmosphere and Oceans*, Wiley, New York, 1978.
19. Veizer, J. and Hoefs, J., *Geochim. Cosmochim. Acta*, 1976, **40**, 1387–1395.
20. Berkner, L. V. and Marshall, L. C., in *The Origins and Evolution of Atmospheres and Oceans* (eds. Brancazio, C. J. and Cameron, A. C. W.), Wiley, New York, 1964.
21. Simkiss, K., *Trans. R. Soc. Edinburgh. (Earth Sci.)*, 1989, 193–199.
22. Chaudhri, R. S. and Kaltia, C. K., in *Conference Volume Geology of Krol basin in Garhwal, Garhwal University*, 1987, pp. 11–37.

ACKNOWLEDGEMENTS. We thank Directors NGRI and WIHG for supporting this programme and Drs. M. Schidlowski (Germany) and V. Divakara Rao (India) for useful discussions and suggestions.

Received 24 November 1993; revised accepted 18 October 1995

Hypoglycaemic and antihyperglycaemic effects of *Aegle marmelos* leaves in rabbits

V. V. Rao, S. K. Dwivedi, D. Swarup and S. R. Sharma

Division of Medicine, Indian Veterinary Research Institute, Izatnagar 243 122, India

Hypoglycaemic and antihyperglycaemic effects of an aqueous extract of *A. marmelos* leaves were studied in normal and alloxanized-rabbits respectively. In normal rabbits, the highest decrease (35.3%) in blood glucose level was recorded with 1 g equivalent dose after 4 h of administration. The hypoglycaemic effect at 12 h was moderate but no effect was observed at 24 h of drug administration. In diabetic rabbits, the extract produced significant ($P < 0.01$) antihyperglycaemic effect within 3 days when given at the dose equivalent to 1 g powder/kg/day.

AEGLE marmelos Linn. (Rutaceae; Bael) is a widely distributed plant in India and the leaf juice of this plant is used orally by Indian Ayurvedic practitioners for the management of diabetes mellitus[1,2]. However, scanty scientific information is available on the hypoglycaemic activity of *A. marmelos* leaves. We report here the role of an aqueous extract of *A. marmelos* leaves as a hypoglycaemic agent in normal and alloxan diabetic rabbits.

Fresh green leaves of *A. marmelos* procured locally were dried in the shade, powdered and boiled with distilled water. After filtration through Whatman filter paper No. 40, the extract was dried by slow heating and continuous stirring. The yield of the extract was 32%.

Adult albino rabbits of either sex and weighing 1.5–2 kg were used in these experiments. The animals were maintained on standard feeding and managemental condition[3].

Hypoglycaemic effect of the extract was evaluated in 24 fasted albino rabbits divided into four groups of 6. The animals of group-I received 20 ml of distilled water *per os*, while the other three were treated orally with an aqueous extract equivalent to 0.5, 1.0 and 1.5 g powder/kg respectively. Blood glucose was estimated[4] at 0, 4, 12 and 24 h following administration of the test drug and the percentage glycaemia variation was calculated for each group. A comparison was also made between 0, 4, 12 and 24 h values of different groups.

Antihyperglycaemic effect was evaluated in alloxan-diabetic rabbits[5] divided into three groups of 6. Group-I received extract equivalent to 1 g powder/kg/day, group-II was given reference drug (Phenformin @ 125 mg/rabbit) while group-III (untreated control) received 20 ml of distilled water. Treatment schedule was continued for 10 consecutive days. The diabetic blood glucose value (day 0) was compared with post-treatment values on days 3, 5 and 10 in all the groups. Comparison was also made between glucose values of different groups at different time intervals. The data were analysed statistically using analysis of variance and paired '*t*' test[6].

Result of the effect of aqueous extract of *A. marmelos* leaves on blood glucose level of normal rabbits is presented in Table 1. The test drug in all the three doses produced significant ($P < 0.01$) hypoglycaemic effect after 4 h of administration. However, it was more marked in animals receiving aqueous extract equivalent to 1 g powdered leaves (group-III). The hypoglycaemic effect was moderate with all the three doses at 12 h

Table 1. Effect of an aqueous extract of *A. marmelos* leaves on blood glucose level of normal rabbits (mean ± SE)

Treatment	Dose equivalent to g powder/kg	Blood glucose levels (mg/100 ml)			
			Post-treatment (h)		
		0	4	12	24
Distilled water	20 ml	97.1 ± 0.8^{A}	96.0 ± 0.9^{A} (1.1)	93.3 ± 0.7^{A} (3.9)	89.7 ± 0.6^{A} (7.6)
Extract	0.5	95.8 ± 1.0^{A}	74.3 ± 1.2^{B} (22.4)	85.5 ± 1.0^{B} (10.7)	89.8 ± 0.4^{A} (6.2)
Extract	1	96.2 ± 1.3^{A}	62.2 ± 0.8^{C} (35.3)	81.4 ± 0.7^{C} (15.4)	88.0 ± 1.5^{A} (6.2)
Extract	1.5	94.4 ± 1.2^{A}	73.1 ± 1.4^{B} (22.6)	78.8 ± 0.8^{C} (16.5)	86.4 ± 1.6^{A} (8.4)

Figures in parentheses indicate % reduction in glucose levels over 0 h.
Capital letters indicate between group comparison.
Means having different letters are significant ($P < 0.01$) but same letters are non-significant.

Table 2. Effect of an aqueous extract of *A. marmelos* leaves and phenformin on blood glucose levels of diabetic rabbits (mean ± SE)

		Blood glucose levels (mg/100 ml)			
		Diabetic level	Post-treatment (days)		
Treatment	Dose	0	3	5	10
Extract	Equivalent to 1 g powder/kg	263.1 ± 4.7^{A}	201.6* ± 5.3^{C} (–23.4)	184.7* ± 5.9^{B} (–29.8)	129.4* ± 3.4^{C} (–50.8)
Phenformin	@ 125 mg per rabbit	278.7 ± 6.0^{A}	239.9* ± 8.2^{B} (–13.9)	186.1* ± 5.5^{B} (–33.2)	114.6* ± 2.1^{B} (–58.9)
Distilled water (untreated control)	20 ml	276.1 ± 4.8^{A}	292.7* ± 4.0^{A} (+ 6.0)	314.2* ± 4.2^{A} (+ 13.8)	333.8* ± 4.0^{A} (+ 20.2)

Figures in parentheses indicate % reduction (–), or increase (+) over diabetic level.
*Significantly different from diabetic level ($P < 0.01$).
Capital letters indicate between group comparison.
Means having different letters are significant ($P < 0.01$) but same letters are non-significant.

and no significant difference in blood glucose level between treated and control group was observed at 24 h.

Table 2 summarizes the effect of an aqueous extract of *A. marmelos* and phenformin on blood glucose level of diabetic rabbits. Both phenformin and aqueous extract of *A. marmelos* were found to have significant ($P < 0.01$) antihyperglycaemic effect after day 3 of treatment. It is interesting to note that on day 3, aqueous extract of *A. marmelos* had comparatively higher antihyperglycaemic effect, while phenoformin was found more effective in reducing blood glucose level on day 10. In untreated diabetic rabbits there was a significant ($P < 0.01$) rise in blood glucose levels from day 3 onward.

These findings indicated that an aqueous extract of *A. marmelos* had a significant hypoglycaemic effect in normal rabbits which lasted for 12 h and the effect was dose-dependent up to 1 g equivalent of powder. However, the response decreased at 1.5 g/kg dose. Such a phenomenon of less hypoglycaemic response at higher doses is not uncommon with indigenous plants and has been observed with *Vinca rosea*[7] and *Cinnamomum tamala*[8]. The extract also produced marked decrease in blood glucose level in diabetic animals. The possible mechanism of the hypoglycaemic effect of *A. marmelos* in normal and diabetic rabbits is not clear. It may be supposed that in normoglycaemic rabbits, the extract exerted hypoglycaemic effect similar to phenformin by stimulating β-cells to release insulin. However, in diabetic rabbits the extract could not have acted by stimulating the β-cells as the alloxan treatment causes permanent destruction of the β-cells. *Ocimum sanctum* leaves[9] and seeds of *Cajanus cajan*[10] have been reported to produce an antidiabetic effect by inhibiting absorption of glucose from the intestine. The possibility of reduced absorption of glucose from the intestine after *A. marmelos* administration cannot be ruled out. A similar explanation has been put forward for antidiabetic activity of *Musa sapientum*[3] and *Caesalpinia bonducella*[5]. Aqueous extract of *A. marmelos* was, therefore, found to have antihyperglycaemic activity in normal as well as alloxanized rabbits, but the precise mechanism whereby this effect is mediated warranted further investigations.

1. Alam, M. M., Siddiqui, M. B. and Hussain, W., *Fitoterapia*, 1990, **61**, 240–242.
2. Prakash, V., *Jeevaniya Shishir*, 1992, **4**, 10.
3. Rao, V. V., Dwivedi, S. K. and Swarup, D., *Fitoterapia*, 1994, **65**, 65–67.
4. Fings, C. S., Tatlif, C. R. and Dunn, R. T., in *Clinical Chemistry* (eds Toro, C. and Ackerman, P. G.), Little Browning and Company, Boston, 1970.
5. Rao, V. V., Dwivedi, S. K. and Swarup, D., *Fitoterapia*, 1994, **65**, 245–247.
6. Snedecor, G. W. and Cochran, W. G., *Statistical Methods*, 6th edn, IBH and Oxford, New Delhi, 1967.
7. Chattopadhyay, R. R., Sarkar, S. K., Ganguly, S., Banerjee, R. N. and Basu, T. K., *Indian J. Physiol. Pharmacol.*, 1991, **35**, 145–151.
8. Sharma, S. R., Dwivedi, S. K. and Swarup, D., *Indian J. Exp. Biol.*, 1995, (in press).
9. Janabai, G., Sugunthi, B. and Meera, G., *Indian J. Nutr. Dietet.*, 1987, **24**, 337–341.
10. Giri, J., Sugunthi, B. and Kowsalya, S., *Indian J. Nutr. Dietet.*, 1986, **23**, 82–87.

Received 11 July 1995; revised accepted 12 October 1995

Do dholes (*Cuon alpinus*) live in packs in response to competition with or predation by large cats?

Arun B. Venkataraman

Department of Entomology, University of Agricultural Sciences, Bangalore 560 065, India and
Indian Institute of Science Field Station, Centre for Ecological Sciences, Masinagudi 643 223, India
*Centre for Ecological Sciences, Indian Institute of Science, Bangalore 560 012, India

Interactions between carnivores during the defence of kills may be one reason why certain carnivores live in groups. This is especially true of lions, hyaenas and the African wild dog. The dhole or the Asiatic wild dog, primarily a pack living animal, has been observed to regularly interact with both tigers and leopards. Such interactions have taken place over kills and otherwise. In this report, five such interactions are described. It was found that the pack's behaviour of surrounding bushes and trees on which the cat was confined precluded immediate escape. The presence of sentinels, while the pack was resting, warned the pack of the presence of a big cat and the pack grouped when a big cat appeared. Costs to both individuals within the dhole packs and the cats involved in the encounters were found to be slight. The reasons for such potentially costly encounters could be competition for finite food resources or thwarting predation. Dholes have a significant diet overlap with both leopards and tigers and aggressively encounter with leopards but not with tigers. Differences between diet overlaps may not be the basis behind the differences in aggression. It is more likely that, the small size of leopards and the fact that they predate more often on dholes, cause dhole packs to be more aggressive to leopards than to tigers. The size of carnivore groups may thus pose an advantage during competitive interactions among carnivore species.

AGONISTIC interactions between carnivore species have been observed during the defence of kills. Lions (*Panthera leo*) regularly defend their kills from the spotted hyaena (*Crocuta crocuta*)[1]. One of the adaptive forces moulding the formation of packs in the African wild dog (*Lycaon pictus*) may be the advantage in numbers for the defence of kills from other species such as the spotted hyaena and for appropriation of kills from other carnivores. In the Selous National Park, Tanzania, wild dogs co-operated in the defence of kills from spotted hyaenas which managed to appropriate only 2% of wild dog kills (Creel and Creel, unpublished manuscript). In contrast, in the Serengeti National Park where hyaenas live in larger groups, the duration that wild dogs retained their kills depended on the ratio of individuals of the two species present[2]. Selous wild dogs often appropriated kills from other carnivores. Very pertinent to this report were the observations that leopards were driven up trees on two occasions and on one of these occasions a kill was appropriated (Creel and Creel, unpublished manuscript). Hyaenas and lions were also occasionally driven off kills which were later appropriated (Creel and Creel, unpublished manuscript).

Accounts of interactions between dhole or Asiatic wild dog packs (*Cuon alpinus*) (for a description of dhole biology, see refs. 3 and 4) and other major carnivore species do occur, with most of the accounts concerning those with the tiger (*Panthera tigris*) and the leopard (*Panthera pardus*)[3,5–13]. The outcome of these encounters though more often favouring the dhole packs involved (when tigers or leopards were killed or driven off), did on occasions result in members of dhole packs getting wounded or killed.

Here I report five separate instances of a dhole pack interacting with four leopards and one tiger (Table 1). One encounter involved one of our study packs and four, the other. The two dhole packs involved in the encounters have been monitored since 1988 and have their home ranges within the Dr J. Jayalalitha sanctuary (formerly Mudumalai sanctuary), Nilgiris District, Tamil Nadu, India[4]. Corroborating the above observed encounters are three other encounters (dhole packs interacting with two leopards and one tiger) which have been reported to the author by other observers. Based on this evidence I feel that such encounters are at the least, not infrequent and have a strong basis behind them. This paper therefore describes the above five encounters and attempts to look at two hypotheses as to why such encounters take place. The first hypothesis is that dhole packs aggressively encounter with leopards and prefer not to with tigers, as the former's prey may significantly overlap with that of dholes while the latter's does not. Coupled with overlap in prey are the relative differences in body weights between the two cats in question and a pack of dholes. Leopards weigh 2–3 times less than tigers do and are an easier adversary for dholes to take on[14]. Even if tigers' prey overlap significantly with those of dholes, the greater body weight differences may result in a high cost of interaction causing the interacting dhole pack to avoid tigers. The second hypothesis is that leopards may prey on single dholes and the aggressive encounters occurring within common hunting areas may drive away leopards, at least temporarily, thwarting sneak attacks on single dholes separated from the pack while hunting.

Table 1 summarizes the details of the above encounters. In all cases where dhole packs encountered leopards, a considerable amount of effort was deliberately put into

*For correspondence

Table 1. Details of dhole–large cat interactions

Date	Cat	Estimated wt. of cat (kg)	No. of adults	No. of pups ≤ 1 yr	Context	Behaviour of cat	Behaviour of pack	Time spent in interaction (min)	Consequence
21.8.93	Leopard	50	8	6	None	Climbed 4 separate trees	Entire pack surrounded all 4 trees on which cat climbed and initially prevented cat from climbing down	50	Leopard escaped without wounds or wounding any dhole
9.1.94	Leopard	80	5	1 (1 yr 1 month)	Den	Climbed bush	Entire pack surrounded bush on which cat climbed and prevented cat from climbing down	96	Leopard escaped without wounds, inflicted scratches on 2 dholes
23.5.94	Leopard	50	11	6	None	Climbed tree	Entire pack surrounded tree	20	Leopard escaped without wounds or wounding dholes
16.9.94	Tiger	> 200	10	8	None	Initially approached pack and then moved away	Single dhole gave alarm calls after seeing tiger. Single dhole confronted tiger and then moved away. Pack moved away and bunched up.	5	Mutual avoidance resulting in no wounds to tiger or dholes
18.1.95	Leopard	50	8	2	Chital kill	Leopard driven from its kill and then driven up tree	Dholes kept leopard up tree after hunting sambar unsuccessfully and consumed entire carcass	30	Leopard left after dholes finished eating carcass and left

harassing the carnivore. The encounter with the tiger suggests, in contrast to the literature, that dholes may prefer not to engage directly with a larger adversary. An earlier account, reported by an observer to the author, which occurred during the study period, confirmed this speculation. A pack of six animals was enough to engage a leopard for 1 h 35 min with a small cost to pack members. Ten dholes were sufficient to deter a tiger from confronting with the pack. Nine dholes could appropriate a kill from a leopard, after hunting unsuccessfully for at least the last 24 h, thereby providing an alternative source of food.

The first hypothesis tested is that such interactions arise over competition for finite food resources. Out of 12 species the three carnivores feed on[15], we have chosen 5 for subsequent analysis. The hair of 5 species have been identified in the scats of the three carnivores, collected from the study area (Arumugam, unpublished). Because of the lack of density estimates, data from black naped hair (*Lepus nigricollis*) has been deleted from any analysis involving density estimates. Observed values of the 4 prey species for each predator are the occurrences of hair of each prey species among the total scats collected from that predator. In a chi-square test for independence, expected values are obtained by assuming that hair occurs among scat in accordance to the number of scats produced from a kill of each prey species and the densities of the 4 prey species within the home ranges of the 2 study dhole packs[15]. From a 3×4 contingency table (Table 2), with the 4 prey species as rows and the three carnivores as columns, we obtain a chi-square value of 1158 (d.f. = 6, $P < 0.01$). All three carnivores kill far more chital and sambar than expected but less gaur and cattle which gives rise to the highly significant result. A niche overlap index was also calculated[16]. Black-naped hare and another class called 'others'[4] has been included in the analysis. The values obtained are dhole–leopard: 0.922, dhole–tiger: 0.915, leopard–tiger: 0.939. This analysis indicates that there is considerable over lap between the diets of each of the three carnivores and the overlap is similar for all carnivore pairs. Therefore, if prey overlap is the reason behind the dholes aggression towards carnivores, they should be equally aggressive to leopards and tigers.

The other hypothesis tested is that leopards may predate on single dholes and such encounters deter them from doing so, by driving them away, at least temporarily, from common hunting areas. Predation on dholes by leopards, though not observed during the study period (October 1990–March 1995), has been reported in the literature[11]. Dhole hair occurred in 2% of leopard scat (as compared to 0.6% of tiger scat)[15]. Dholes while resting, always have a sentinel and are alerted by alarm

Table 2. Observed and expected numbers of occurrences of hair of 4 prey species found in scat of dhole, leopard and tiger

	Dhole		Leopard		Tiger	
Prey	Obs.	Exp.	Obs.	Exp.	Obs.	Exp.
Chital (*Axis axis*)	501 (172.40)	280.93	117 (16.60)	80.45	43 (6.23)	29.45
Sambar (*Cervus unicolor*)	243 (344.54)	78.52	76 (110.93)	24.19	48 (101.53)	12.45
Cattle	136 (128.56)	347.30	82 (7.54)	110.93	19 (11.36)	40.43
Gaur (*Bos gaurus*)	0 (173.25)	173.25	1 (58.44)	60.42	1 (26.70)	28.67

Total chi-square value = 1158.09
Densities (in km^2): Chital = 25.03, Sambar = 6.61, Cattle = 23.39, Gaur = 14.38. Data from ref. 20.
Chi-square values for each cell entry, in brackets.

calls of other mammals such as the common langur (*Presbytis entellus*), chital (*Axis axis*) and sambar (*Cervus unicolor*) (Venkataraman, unpublished). Furthermore, in the second encounter described, the leopard was in the proximity of a den containing young pups. It is very likely that harassment may deter sneak attacks on single dholes and pups. Dholes often separate during hunting and have elaborate acoustic communication for reassembling[3]. As mentioned earlier, packs have vigilant members who may thus deter attacks on the group.

In the Bandipur National Park, Karnataka, India, leopards and dholes competed for similar-sized prey but competition was reduced by the leopard's ability to climb and cache prey in trees. Moreover dholes killed more fawns and aged chital, while leopards killed chital of all ages. Tigers may coexist with dholes because of several factors such as spatio-temporal use of habitat and sex and size of prey[17]. Leopards and dholes killed prey of the same weight class while tigers killed prey of a higher weight class. Dholes, however, had a preference for male chital, while leopards did not. Tiger predation was biased towards adult males in chital, sambar, wild pig and towards young gaur[15]. Even though a significant diet overlap has been reported in this paper, competition may be reduced by niche partitioning based on age and sex differences of the prey species and spatio-temporal use of the habitat and prey. This may still cause a reduction and consequent difference in the overlap between the diets of dholes, leopards and tigers. However it is more likely that because leopards are smaller in size and predate more frequently on dholes, they are more aggressively dealt with than tigers are. Even though the ultimate reasons for such confrontations cannot be elucidated with certainty, it is quite evident that a pack could successfully accomplish what a single dhole or even a pair may not. Pack members successfully kept leopards on trees and bushes by surrounding them. A pack could obtain food by driving a leopard away from its kill. A single dhole detected a tiger and warned the pack. The pack bunched up invoking ideas of the selfish herd paradigm[18]. The tiger did not directly confront the pack avoiding a major cost to the pack.

Speculations about the evolution of group size in carnivores have hitherto concentrated primarily on the benefits of numbers while hunting[19]. This report brings to light another advantage of numbers in geographical regions where conflicts between major carnivores may be a definite selective force for the evolution of groups.

1. Lamprecht, J., *Z. Tierpsychol.*, 1978, **46**, 337–343.
2. Fanshawe, J. H. and Fitzgibbon, C. D., *Anim. Behav.*, 1993, **45**, 479–490.
3. Johnsingh, A. J. T., *J. Zool. (London)*, 1982, **198**, 443–463.
4. Venkataraman, A. B., Arumugam, R. and Sukumar, R., *J. Zool. (London)*, in press.
5. Arjansingh, B., *Tiger Haven*, Macmillan, London., 1973, p. 232.
6. Burton, R. W., *J. Bombay. Nat. Hist. Soc.*, 1925, **30**, 910–911.
7. Clive, C. M. J., *J. Bombay. Nat. Hist. Soc.*, 1928, **32**, 590.
8. Connell, W., *J. Bombay. Nat. Hist. Soc.*, 1944, **44**, 468–470.
9. Khajuria, H., *J. Bombay. Nat. Hist. Soc.*, 1963, **60**, 448–449.
10. Morris, R. C., *J. Bombay. Nat. Hist. Soc.*, 1925, **30**, 218–219.
11. Morris, R. C., *J. Bombay. Nat. Hist. Soc.*, 1925, **30**, 691–693.
12. Morris, R. C., *J. Bombay. Nat. Hist. Soc.*, 1933a, **36**, 744–745.
13. Morris, R. C., *J. Bombay. Nat. Hist. Soc.*, 1933b, **36**, 744–745.
14. Karanth, K. U., in *Tigers of the World* (eds Tilson, R. L. and Seal, U. S.), Noyes Publications, Park Ridge, 1987, pp. 118–133.
15. Karanth, K. U. and Sunquist, M. E., *J. Anim. Ecol.*, in press.
16. Pianka, E. R., *Annu. Rev. Eco. Syst.*, 1973, **4**, 53–74.
17. Johnsingh, A. J. T., *J. Bombay. Nat. Hist. Soc.*, 1983, **80**, 1–57.
18. Hamilton, W. D., *J. Theor. Biol.*, 1971, **33**, 295–311.
19. Packer, C. and Ruttan, L., *Am. Nat.*, **132**, 159–198.
20. Varman, K. S. and Sukumar, R., *J. Biosci.*, 1995, **20**, 273–287.

ACKNOWLEDGEMENTS. This work was supported by the Department of Science and Technology, Government of India, under the Young Scientists Programme. I thank Drs N. V. Joshi, Anindya Sinha, R. Sukumar and two anonymous referees for their critical comments while preparing the manuscript. I would also like to thank K. Kunte, M. Kethan and M. Siddhan for excellent assistance in the field.

Received 14 March 1995; revised accepted 9 October 1995

Estimates of global methane production from rice paddies based on substrate requirement

Sabyasachi Ghosh, M. C. Jain and Suresh K. Sinha

Division of Environmental Sciences and Water Technology Centre, Indian Agricultural Research Institute, New Delhi 110 012, India

Estimates of methane emission from rice paddies are highly variable and hence debatable. Most of the estimates were made without consideration of the substrate requirements. A mole of glucose should normally produce 2 moles of methane yielding 178 mg methane from 1 g of glucose. In our incubation experiment only 10–11% of this potential efficiency was realized. Estimates of biomass of globally produced rice were made from the yield of rice and thus of the potential leaching of the organic substances which gives an upper limit of 13 Tg global methane production from rice paddies.

RICE paddies have been identified as an important source of methane[1]. This was based on some experiments which did not simulate rice-growing conditions of Asia where 92% of the world's rice is grown. The experiments in Europe and US had high organic matter and nitrogen in the soil (200 kg N ha^{-1}) and continuous standing water with an yield of 7000 kg ha^{-1} as against Asian soils with low organic matter and low fertilizer input, which varies from none to 150 kg N ha^{-1} with an yield average of hardly 2000 kg ha^{-1}. In most of the experiments the emphasis was on developing suitable techniques of sampling and gas chromatography. However, the important observations on diurnal variation in methane emission, and the maximum flux of methane coinciding with the maximum biomass of the crop were not utilized for understanding the limits of methane production. The microbial and biochemical basis of methane production has been known[2], according to which a source of carbon is essential for methane production. Here we report the estimates of methane emission from rice paddies on the basis of requirement of carbon source.

Rice-paddy soils are the source of methane, for which there ought to be a source of carbon. This could be either organic carbon or constant supply of carbon source (carbohydrate or organic acids) from the crop itself[3]. Since soil organic carbon is apparently not a major source of methane because about no methane is produced from soils without a rice crop, hence, it was assumed that a carbohydrate such as glucose would simulate rice-paddy soil for methane production.

Soil samples were taken at random from a depth of 0–15 cm from a plot of 500 m^2 after the rice harvest. The samples were air-dried and sieved through 2 mm mesh screen. The soil samples were incubated at 30± 1°C and 35± 1°C temperature in conical flasks of 1 l capacity each. This range of temperature was chosen because at this temperature rice cultivation in India takes place and methanogenic activity is maximum. Provisions were made in the flask for measurements of pH and Eh, collection of gas samples for methane estimation and collection of liquid samples for estimation of acetate content.

Two hundred and fifty grams of air-dried soil was taken in each flask and distilled water was added to cover the soil with a 5 cm deep layer of water. Nitrogen gas (99.998%) was flushed every 3 days for 5 min to ensure complete anaerobic conditions. Two treatments, viz. control (no glucose) and 1 g glucose, were given to the soil (250 g) in different incubation flasks in three replicates.

Gas samples were collected using 1 ml gas-tight syringe through a three-way valve at a regular interval of 24 h for 30 days. Supernatant water samples were also drawn from the side port of the flask to estimate acetate content. Volume of samples was replenished by equal amount of water.

Gas samples were analysed for methane on a Packard Gas Chromatograph (Model 438) using flame ionization detector (FID)[4]. Stainless steel column packed with molecular sieve 5 Å 60–80 mesh at a temperature of 80°C was used for methane analysis. The injector and detector were kept at 100°C. Nitrogen at a rate of 15 ml min^{-1} and H_2 at a rate of 25 ml min^{-1} were passed through as carrier and fuel gases respectively. The equipment was calibrated and frequently checked against the standard gas samples.

Five ml sample of supernatant aqueous phase above the soil was obtained using a syringe through a side port and was passed through Whatman No. 1 filter paper. The acetate concentration was measured using an Ion Chromatograph (Dionex, Dionex Corporation, USA) fitted with a suppressed conductivity detector and integrator. The column used was 4S 4A-5C. The eluent was a mixture of 1.7 mM sodium bicarbonate and 1.8 mM sodium carbonate in the ratio of 1 : 1 with a flow rate 2 ml min^{-1}. The pump pressure was 1160 psi. Samples were injected through a loop of 25 µl automatically by a pneumatic pressure-operated injector and the excess volume was flushed out. The anions were separated in the column and detected by the conductivity detector and the chromatograms were recorded by the integrator. The acetate concentration in the samples was determined with reference to a standard solution of 30 ppm acetate prepared from ammonium acetate (AR grade).

The pH was measured using the combined electrode in the soil solution through the side port of the flask using digital pH meter, Century-model CP 901. Eh (redox potential) was recorded using the same pH cum mV meter. A platinum micro electrode prepared in the

laboratory was installed permanently in the flask and a calomel electrode through side port served as a reference electrode.

Methane production from control (absence of glucose) at 30°C showed a steep rise from day 7 onwards and reached a maximum value of 0.07 $\mu g\ g^{-1}\ day^{-1}$ by day 18. The production then declined to 0.01 $\mu g\ g^{-1}\ day^{-1}$. The pH during the experimental period varied between 7.1 and 7.9. However, Eh reached – 157 mV by day 13 and then remained between – 183 and – 234 mV from day 17 onward (Figure 1). The correlation coefficient between methane emission and Eh and pH independently was – 0.90 and 0.16 respectively (Table 1).

The incubation of the soil with 1 g glucose gave an emission of 8.88 $\mu g\ g^{-1}\ day^{-1}$ on day 15, being 130 times higher than the control (Figure 1). The Eh also declined sharply to – 480 mV by day 10 and reached – 498 mV by day 25. Initially, the pH was reduced to 6.2 but with the decline in methane production it increased and reached a value of 7.9. The correlation coefficient between emission and Eh was – 0.82, and between methane emission and pH it was – 0.14. Table 1 gives the correlation coefficient between methane emission with Eh and pH. At 35°C, the maximum production of methane (9.03 $\mu g\ g^{-1}\ day^{-1}$) was observed on day 10. The Eh of the system declined below –200 mV by day 12 and reached –410 mV on day 14. The production of methane declined to 0.154 $\mu g\ g^{-1}\ day^{-1}$ by day 30.

Acetate accumulation started by day 5 and day 6 when the Eh was still above –100 mV. In the treatments of 1 g glucose at 30 and 35°C acetate accumulation preceded methane emission. Figure 2 shows this relationship at 35°C.

Methane production from rice paddies has been estimated considering microbiological processes and also the potential requirements of substrate for methane emission[3]. Since plants do release organic substances, such as, sugars and organic acids, these substances would serve as substrates for methane production and other microbial activities such as growth, multiplication and production of metabolites. Assuming that eventually glucose is the substrate, then 32 g methane would be produced from 180 g glucose if there was no reduction of CO_2 to methane or 178 mg methane would be produced for 1 g of glucose. In our experiment the maximum amount of methane production was 17.49 mg and 19.16 mg from 1 g glucose at 30 and 35°C respectively. This means that the substrate utilization efficiency was only 10% and 11% at 30 and 30°C respectively.

The cause of such low efficiency of the conversion of glucose to methane may be due to the following reasons: a) A part of the glucose may be converted to carbon dioxide or got bound in clay organic matter complexes, thus reducing efficiency; b) Inhibitory nature of some of the end product in the process of methane formation may reduce efficiency; c) Methane oxidation in the upper layers of the soil by methanotrophic bacteria may reduce efficiency.

These data can now be used for assessing the carbo-

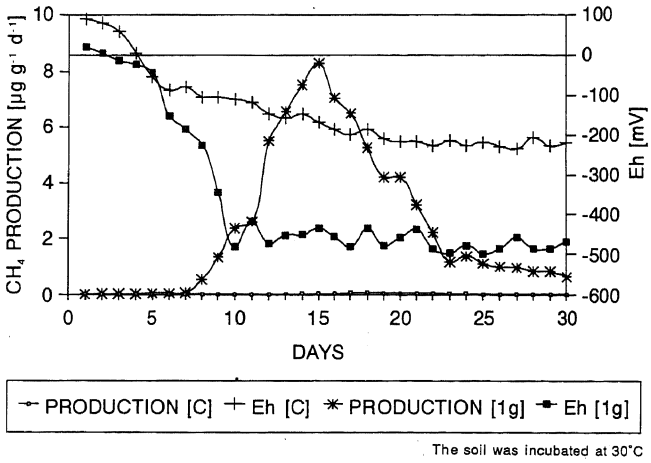


Figure 1. Time-course of redox and methane production with or without glucose amendment.

Table 1. Correlation coefficients for the relationship of methane emission with Eh and pH in soils with and without glucose amendment

Incubation temperature	Treatments	Eh	pH
At 30°C	Control	–0.90*	0.16
	1 g glucose	–0.82*	–0.14
At 35°C	Control	–0.91*	0.04
	1 g glucose	–0.90*	0.03

*Significant at 1%.

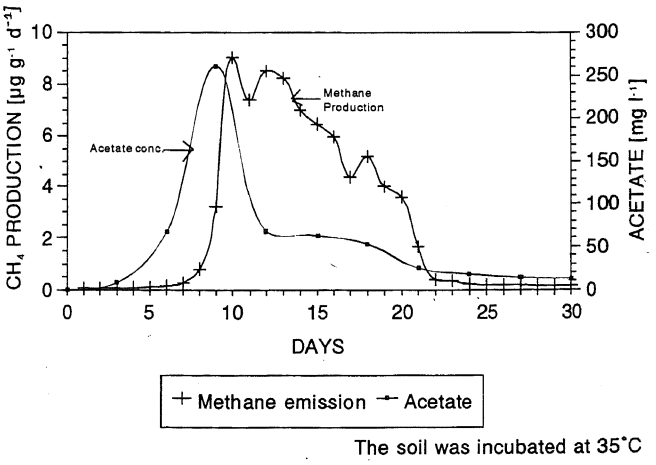


Figure 2. Effect of 1 g glucose as a substrate on methane emission and acetate accumulation.

hydrate requirements for methane production. According to the IPCC (1992), the annual average production of methane from rice paddies is 110 Tg. If the efficiency, as experimentally observed, is taken into consideration, it would require 1100 million tonnes (10% efficiency) or 1000 million tonnes (11% efficiency) of glucose or glucose equivalent for producing 110 million tonnes of methane. Assuming the harvest index of 40%, with an annual rice paddy production of 527 million tonnes, the maximum biomass would be 1300 million tonnes. All estimates made so far suggest that plants can release organic substance up to 10% of their weight through roots[5]. Thus, a maximum amount of 130 million tonnes may be available as substrate for methane production. From this amout only 11 to 13 million tonnes of methane can be produced. Thus an upper limit of 13 Tg methane production from rice paddies can be expected. Any assessment ignoring biological aspects of the mechanism of methane production is an artifact of methodology.

1. *Climate Change—The IPCC Scientific Assessment* (eds Houghton, J. T., Jenkins, G. J. and Ephraums, J. J.), Cambridge University Press, 1990.
2. Wassman, R., Papen, H. and Rennenberg, R., *Chemosphere*, 1993, **26**, 201–217.
3. Sinha, S. K., *Curr. Sci.*, 1994, **68**, 643–646.
4. Cicerone, R. J. and Shetter, J. D., *J. Geophys. Res.*, 1981, **86**, 7203–7209.
5. Tukey, H. B., *Annu. Rev. Plant Physiol.*, 1970, **21**, 305–324.

Received 7 August 1995; revised accepted 12 October 1995

Bacillus thuringiensis, a biocontrol agent for major tea pests

N. Unnamalai and Vaithilingam Sekar

School of Biotechnology, Madurai Kamaraj University, Madurai 625 021, India

A preliminary screening was carried out to determine the efficacy of crystal proteins of nine strains of *Bacillus thuringiensis* to control two lepidopteran pests of tea plants. Among these strains *B. thuringiensis* var. *kurstaki* HD1 and HD73 were highly toxic, *B. thuringiensis* var. *thuringiensis* HD2 was moderately toxic, *B. thuringiensis* var. *aizawai* HD133 showed low toxicity, while the other strains showed no toxicity towards these pests. The LC_{50} values of the three toxic strains and that of the purified crystal protein of the highly toxic strain HD1 were determined by laboratory bioassays.

TEA is an important plantation crop of India and is affected by many pests. The two most economically important lepidopteran insect pests of tea are the leaf-feeding caterpillars, *Caloptilia theivora* (leaf roller) and *Cydia leucostoma* (flushworm), belonging to the family Gracillariidae and Olethreutidae respectively. They cause considerable damage to young leaves of the tea plant by destroying the apical bud and tender leaves, thereby arresting the shoot growth.

The choice of conventional insecticides to be used on this crop is limited because of concerns related to consumer protection and environmental safety. Hence the search for effective biopesticides for the management of tea pests becomes appropriate. We undertook a study on the feasibility of the use of *Bacillus thuringiensis*, a well-known entomopathogen, as a biopesticide in tea plantations.

We carried out a preliminary screening for the efficacy of the crystal proteins of various *B. thuringiensis* strains belonging to class I which are known to be toxic to lepidopterous insects (Table 1). These strains were cultured on nutrient agar plates (peptone, 5 g; NaCl, 5 g; yeast extract, 1.5 g; beef extract, 1.5 g; distilled water, 1000 ml; pH 7.4) for four days at 30°C. One plate was used for each strain. The resultant spore–crystal mixtures were scraped off the plates, washed once in 0.5 M NaCl, followed by three washings in sterile water and resuspended in 1 ml of sterile water containing a protease inhibitor phenylmethylsulphonyl fluoride, at a concentration of 1 mM. These crude preparations were stored at –20°C until further use.

The pure crystal protein preparation was made from the spore–crystal mixture of HD1 (ref. 1). The purity of the crystals was 95% as determined by light microscopy. The level of purity of crystals was also confirmed by SDS–PAGE analysis[2]. Protein concentration of the spore–crystal mixtures of all nine strains and the purified crystal fraction of the strain HD1 were estimated by the dye-binding method[3] after solubilizing the crystal inclusions at 37°C in 10 mM NaOH (pH 10.5) containing 25 mM dithiothretol.

Prior to setting up bioassays, the surface area of the leaves was measured. Preliminary bioassays were conducted using crude spore–crystal samples obtained from the nine strains of *B. thuringiensis* towards larvae of flushworm and leaf rollers. Assays were done at 28°C

Table 1. Various strains of *B. thuringiensis* used in the study

Strain	Serotype
B. thuringiensis var. *kurstaki* strain HD1	3a 3b
B. thuringiensis var. *kurstaki* strain HD73	3a 3b
B. thuringiensis var. *thuringiensis* strain HD2	1
B. thuringiensis var. *aizawai* strain HD133	7
B. thuringiensis var. *indiana* strain HD521	15
B. thuringiensis var. *dakota* strain HD551	16
B. thuringiensis var. *pakistani* strain HD395	13
B. thuringiensis var. *morrisoni* strain HD12	8a 8b
B. thuringiensis var. *galleriae* strain HD207	13

Table 2. Bioassays using the spore–crystal mixtures of various *B. thuringiensis* strains[a]

Strain	Percentage mortality after 72 h[b]	
	Flushworm	Leafroller
HD1	70	60
HD73	60	65
HD2	45	45
HD133	30	25
HD521	0	0
HD511	0	0
HD207	0	0
HD12	0	0
HD395	0	0
Control*	0	0

*Water-coated leaves were used as control.
[a]Leaves coated with 200 ng cm^{-2} of crude crystal protein were used.
[b]Average of three replicates.

Table 3. Toxicity of selected crude/purified crystal proteins against flushworm and leafroller

Strain	Flushworm*		Leafroller*	
	LC_{50}** (ng cm^{-2})	95% fiducial limits	LC_{50}** (ng cm^{-2})	95% fiducial limits
HD1	185.36	151.39–226.95	174.97	143.84–212.82
HD73	208.95	171.70–254.28	184.05	150.31–225.35
HD2	221.79	184.39–266.78	296.67	178.60–492.80
Purified HD1	40.97	37.02–45.35	33.93	30.07–38.28

*Larval mortalities scored after 72 h.
**LC_{50} values were calculated from three independent bioassay tests.

by coating fresh individual tea leaves with 200 ng cm^{-2} of protein from crude spore–crystal preparation and placing the 2nd instar larvae over it. In all bioassays, a total of 20 larvae/treatment (5 larvae/leaf) were used. The percentage mortality was scored after 72 h. A control was maintained by coating sterile distilled water on the leaves. All bioassays were repeated three times.

LC_{50} values of crude spore–crystal samples of the three toxic strains were estimated by conducting larval bioassays. Different concentrations of crude samples (100, 150, 200, 250, 300 ng cm^{-2}) and the pure crystal protein of the most toxic strain (20, 30, 50 and 75 ng cm^{-2}) were used in the bioassays. The data were subjected to probit analysis[4].

The results of preliminary bioassays showed that *B. t.* var. *kurstaki* HD1, HD73 and *B. t.* var. *thuringiensis* HD2 were toxic to the pests. *B. t.* var. *aizawai* showed some toxicity, while the other strains showed no toxicity towards these pests (Table 2).

Among these three strains, HD1 and HD73 were most toxic to both flushworms and leafrollers. The LC_{50} value of HD1 towards flushworms was 185.36 ng cm^{-2} with a 95% fiducial limit (FL) of 151.39–226.95 and that of HD73 was 208.95 ng cm^{-2} (FL = 171.70–254.28). The LC_{50} value of HD1 towards the leaf rollers was 174.97 ng cm^{-2} (FL = 143.84–212.82) and that of HD73 was 184.05 ng cm^{-2} (FL = 150.31–225.35) (Table 3). The LC_{50} value of the purified crystal protein of HD1 was 40.97 ng cm^{-2} and 33.93 ng cm^{-2} respectively for flushworms and leafrollers (Table 3). The low LC_{50} values observed for the purified crystal proteins when compared to the crude preparations may be due to the presence of significant amounts of the spore proteins in the crude samples.

Attempts to control pests of tea plant using *B. thuringiensis* were made by Kariya[5] in 1977. Recently, CryIA(a) protein of Fu-2-7 strain was reported to be about half as toxic against smaller tea tortrix as the strain HD1. It is known that *B. thuringiensis* var. *kurstaki* strain HD1 produces several lepidopteran toxic proteins such as CryIA(a), CryIA(b), CryIA(c), CryIIA and CryIIB while, HD73 produces only CryIA(c) protein[7]. Since the LC_{50} values for the strain HD73 (containing CryIA(c) protein only) and HD1 (containing several Cry proteins) were quite similar, it is likely that CryIA(c) may be more toxic. Other CryI proteins such as CryIA(a), CryIA(b) and CryIB produced by the strains HD1 and HD2 may also exhibit some toxicity. Studies to establish the relative efficacy of individual Cry proteins to control leafrollers and flushworms are underway.

1. Murty, M. G., Srinivas, G., Bora, R. S. and Sekar, V., *J. Microbiol. Methods*, 1994, **19**, 103–110.
2. Laemmli, U. K., *Nature*, 1970, **227**, 680–685.
3. Bradford, M. M., *Anal. Biochem.*, 1976, **74**, 248–254.
4. Finney, D. J., in *Probit Analysis*, Cambridge Univ. Press, London, 1971.
5. Kariya, N., *Study of Tea (Japan)*, 1977, **53**, 51–55.
6. Udayasuriyan, V., Nakamura, A., Mori, H., Masaki, H. and Uozumi, T., *Biosci. Biotechnol. Biochem.*, 1994, **58**, 830–835.
7. Hofte, H. and Whiteley, H. R., *Microbiol. Rev.*, 1989, **53**, 242–255.

ACKNOWLEDGEMENTS. We thank Dr L. K. Nakamura, NRRL, USDA, Peoria, Illinois for providing *B. thuringiensis* strains used in the study. This work was supported by a research grant from Parry Agro Industries Ltd, Madras. We also thank Dr J. Muthukrishnan for his help in the statistical analysis and Mr G. Neethirajan for technical assistance.

Received 29 July 1995; revised accepted 17 October 1995

Morphology of zircons in massive and incipient charnockites of southern Kerala: Their bearing on the origin

G. R. Ravindra Kumar and Narayanaswamy
Centre for Earth Science Studies, Akkulam, Trivandrum 695 031, India

Morphological characteristics of zircons in massive and incipient charnockites show contrasting features between the two types. Significantly greater length/breadth ratio in massive charnockite suggests crystallization in igneous environment, whereas zircons of incipient charnockites point to extended transport, eventual rounding and deposition in sedimentary environment. Further, morphological data agree with the petrological and geochemical observation that the incipient charnockites are probably the highgrade (granulite facies) reaction products of metasediments (argillaceous and arenaceous) whereas massive charnockites developed from igneous protoliths.

ZIRCON is an accessory mineral of great interest in many rocks. The relative abundance, colour, elongation ratio, degree of rounding of zircons can give positive expression as to the nature of origin of the host rock. In the past, several attempts have been made to resolve some controversial petrogenetic problems of distinguishing igneous rock types, origin of granite, metasomatized rocks, different igneous intrusions in an igneous complex, ortho/para nature of gneisses[1–5] using the statistical parameters of zircons. Zircon study has been found especially significant in the study of granite.

We have attempted here the morphological study of zircons to address the mode of origin of two types of incipient and massive charnockites. The incipient charnockite is seen as foliation-blurring discordant coarse-grained patches and the massive charnockite occurs as, occasionally banded, large units in southern Kerala. Our interpretation of statistical data is based on the knowledge of field relationships and the observation of several workers[6–8] that zircons retain a large extent of their premetamorphic characteristics through higher grade of metamorphism and can survive crustal melting events and remain unrecrystallized even under granulite facies conditions.

Southern Kerala is essentially a granulite-facies terrain. A major shear zone called the Achenkovil shear zone separates the high-grade terrain into a northern region dominated by garnet- and graphite-free massive charnockites, from the southern region dominated by metasedimentaries. The metasedimentary unit, also known as Kerala khondalite belt (KKB) consists of graphitic garnet–sillimanite–biotite–cordierite gneiss with less abundant calc-silicate and quartzites. Widespread occurrences of incipient charnockites are noted within the supracrustal assemblages of the KKB. Rb–Sr and U–Pb dating of khondalite group suggest the age of khondalite to be 1900–2200 my (ref. 9–11). It has been suggested by Chacko *et al.*[10] that the khondalite belt may be representing Proterozoic fault-bounded basin that received much of its arkosic sediments from granitic high lands and deposited over the pre-metamorphic equivalent of massif charnockites (tonalitic orthogneiss). The incipient charnockite-forming event appears to be a late Proterozoic event (540 my (ref. 9)). Previous morphological study of zircons of khondalites and massive charnockites suggested respectively sedimentary and magmatic origin for the rock types[12]. In this study we have selected an incipient charnockite from Nedumangad (Lat 8°42″; Long 77°02″) and a massive charnockite 2 km east of Vadasserikara (Lat 9°21″; Long 76°50″) to study the morphological characters of zircons and to verify the data against the prevailing views that (a) charnockites in southern Kerala are magmatic[13] and (b) the massive charnockites are igneous and incipient charnockites formed from metasediments[9,10,14].

The massive charnockite was collected from the northern margin of the Achenkovil shear zone and at the northern terminus of the KKB. The rock is medium grained with greenish to greyish black colour. It occurs as a large massive body with no diagnostic structural element in quarry exposures but the weathered portions reveal gneissic layering. The rock is made up of a typical granulite facies assemblage of quartz + alkali feldspar + plagioclase (An_{40}) + hypersthene + hornblende + magnetite and zircons, set in a hypidiomorphic texture. Graphite and garnet are absent.

The incipient charnockite was collected from a fresh active incipient charnockite quarry near Nedumangad. The charnockite is interlayered with graphitic garnet biotite gneiss. Sillimanite-bearing pelitic gneiss (khondalite) occurs few metres away but in the same interlayered sequence. Distinct planar fabric defined by garnet and biotite is present in the charnockite. No cross cutting relation between gneiss and charnockite, which could be assigned to an igneous process was observed. The field relationships between gneiss and charnockite are similar to those seen in incipient charnockite localities of south Kerala[9,14]. The rock is coarser grained with a mineralogy of quartz + K-feldspar + plagioclase (An_{35}), garnet + biotite + hypersthene + graphite + ilmenite and zircon set in a homogeneous granoblastic texture.

Zircon was separated following the method of Murthy *et al.*[15]. The work involved crushing of samples to –60 mesh, washing to eliminate fine powder fraction and heavy mineral separation at 1.4 amperage with forward slope of 15° and side tilt of 20° using isodynamic separator. The nonmagnetic concentrate containing zircon was treated in methylene iodide and clericii solution.

Table 1. Morphological data of zircons

Charnockite	Colour	Euhedral/ terminally rounded	Rounded	Over growths	Metamict	Inclusions
MC 278: Massive (ortho)	Colourless	97.5	2.5	8	1.5	53.5
IC 279: Incipient (para)	Colourless to light brown	75	25	3.5	2	34.5

Table 2. Statistical data on zircons

Charnockite	$\bar{X}$	$\bar{Y}$	N	SX	SY	r	a	θ	σa	Dd
Massive	0.3027	0.1162	200	0.1185	0.0376	0.5345	0.3173	17°36″	0.0189	13.69
Incipient	0.2009	0.1137	200	0.0891	0.0429	0.0457	0.4815	25°46″	0.0340	35.51

Final purification was made by hand picking under stereobinocular microscope. Two hundred unbroken zircons were used for statistical morphological studies (Tables 1 and 2). Statistical computations are based on the measurement of mean length ($\bar{X}$), mean breadth ($\bar{Y}$), standard deviation of length (SX) and breadth (SY), standard error of slope (σa), coefficient of correlation (r) and relative dispersion (Dd).

Zircons generally occur as tiny crystals in very minor quantities in both the rocks. They are translucent and colourless, which are characteristic of granulite-facies varieties[16,17]. Minor occurrence of hyacinth coloured varieties is noticed. Four types of zircons are noted in both the rock types, viz. euhedral, terminally rounded, subrounded and rounded. Globular and rod-like inclusions are fairly common in zircons but overgrowths and zoning are uncommon.

The statistical parameters of zircons are given in Table 2. In massive charnockite, zircons vary in length from 0.11 to 0.8 mm and breadth from 0.05 to 0.29 mm. Up to 73% is terminally rounded and does not show any angle between length and C-axis. 25% of zircon population is made up of euhedral crystals (Figure 1 *a*). Overgrowths are negligible. The size–frequency study, following the method of Larsen and Poldervaart[6] suggests length maxima, depicting polymodal prominent peak at 0.225 mm and breadth maxima depicting a monomodal, prominent peak at 0.1 mm (Figure 2 *a, b*). The elongation curve is sharp monomodal with a maxima at 2.5 mm (Figure 2 *c*).

In incipient charnockite a high percentage of rounded zircons (~ 25%) was noted (Figure 1 *b*). Euhedral and terminally rounded zircons make about 18% and 57% respectively. Overgrowths are seen in about 4% of the zircons. Variable angle between length and C-axis is noted in the rounded zircons.

The data of length (0.08–0.54 mm) and breadth (0.05–

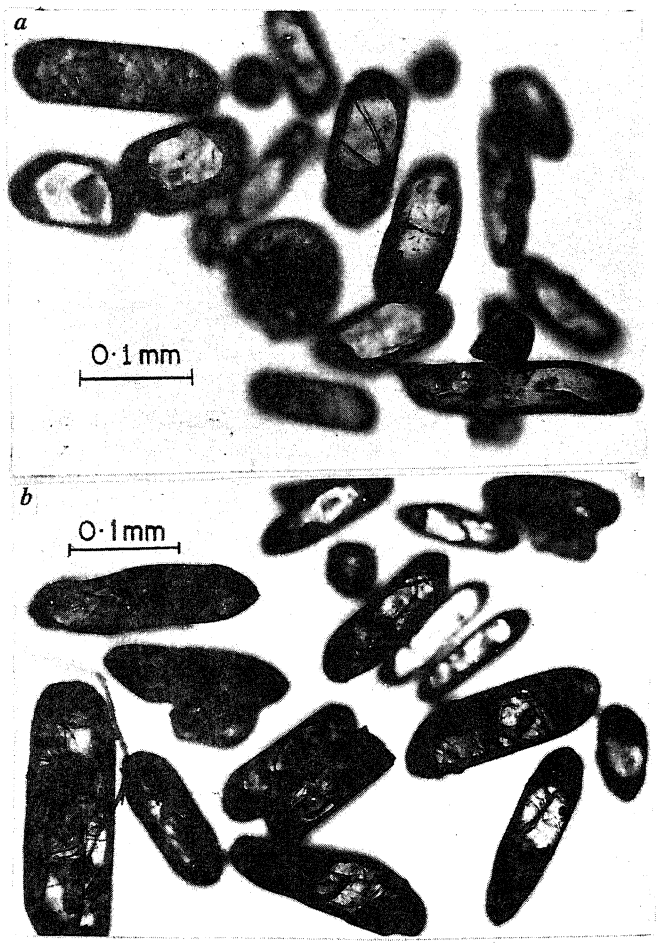


Figure 1. *a*, Zircon crystals from massive charnockite. Note the corroded and terminally rounded margins and euhedral shape. *b*, Zircon crystals in incipient charnockite. Observe that a majority of zircons show rounded and a few show euhedral shapes.

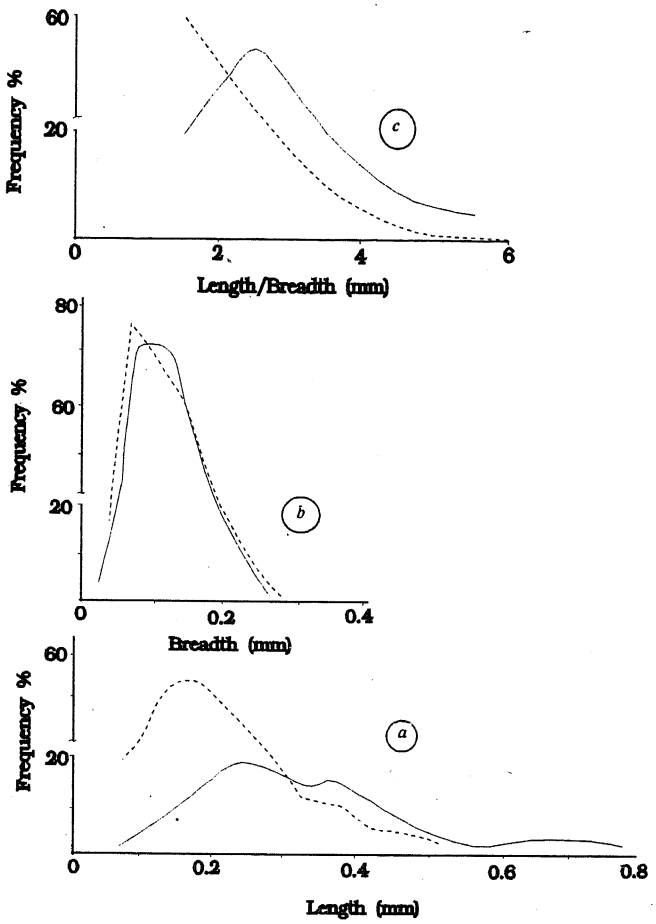


Figure 2. Size frequency data of zircons of massive (unbroken line) and incipient charnockite (broken line) plotted against (*a*) length, (*b*) breadth and (*c*) length/breadth.

0.32 mm) analysed according to Larsen and Poldervaart[6] suggest polymodal length maxima at 0.175 mm, and sharp monomodal prominent peak with a breadth maxima at 0.75 mm (see Figure 2 *a, b*). Elongation curve is sigmoidal (Figure 2 *c*).

A statistical method of treating zircon by reduced major axes (RMA) technique was developed and tested by Larsen and Poldervaart[6]. It is considered to depict the growth trend line of zircon. This trend line also provides a means of evaluating any differences between the rock types. This method follows a simple principle that zircons crystallized from the same magma will have similar physical properties and follow a linear distribution in length–breadth scatter diagram. To test the differences between the incipient and massive charnockites, the reduced major axes of these rocks have been plotted (Figure 3) following the method of Larsen and Poldervaart[6]. The visual comparison of RMA shows significant variation with different slopes for incipient and massive charnockites. The variation in the slope decreasing with increase in elongation of zircons shows that these rock

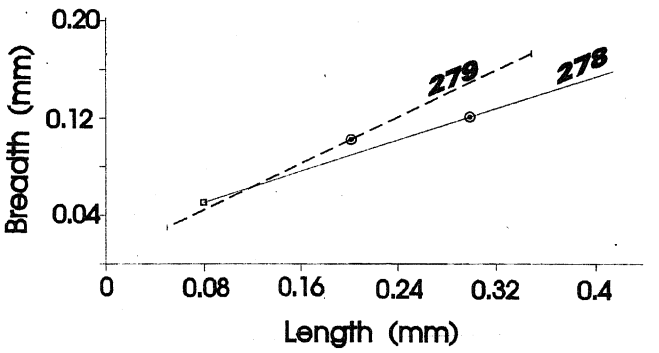


Figure 3. Reduced major axes of length (mm) and breadth (mm) measurement of massive (unbroken line) and incipient charnockite (broken line).

types belong to diverse source region and have entirely different nature of origin. The noncomparable RMA slope position of euhedral and rounded zircons of incipient charnockite suggests dominant detrital nature and the rounding of zircons as due to sedimentary processes. However, similar tests on the zircons of massive charnockite suggest a comparable RMA slope, suggesting the rounding of zircons of these charnockites to be due to corrosion[18].

From the present study it is clear that the characteristics of zircon in massive charnockite are in contrast to those in the incipient charnockites. Higher percentage of euhedral zircons is seen in massive charnockite whereas the zircons of the incipient charnockite are detrital in character and have relatively greater percentage of rounded zircons. This suggests that the precursor rocks of incipient and massive charnockites may have formed in different environments. Terminally rounded nature of zircons ascribed to corrosion, and relatively higher euhedral zircons in massive charnockites favour igneous origin. Moderately well rounded and a few terminally rounded zircons in incipient charnockites suggest diverse source and some of them to be reworked. Further, the zircons from massive charnockite have significantly greater length/breadth ratios than the incipient charnockite zircon varieties. The low length/breadth ratios of zircons of incipient charnockite are consistent with a sedimentary protolith for this rock type. Although less, significant presence of euhedral zircons and terminally rounded zircons in incipient charnockite supports the contention that the immediately adjacent high land areas where massive charnockite are abundant[10] provided the detritus to the depositional basin, whereas the rounded zircons suffered a long distance transport.

1. Alper, A. M. and Poldervaart, A., *Eco. Geol.*, 1957, **52**, 952–971.
2. Murthy, M. V. N., *Zircon*, Misc. Publ. No. 9, Geol. Surv. India, 1968, pp. 1–226.
3. Gupta, L. N., *J. Geol. Soc. India*, 1968, **9**, 124–137.
4. Gupta, L. N. and Elsdon, R., *Ir. J. Earth. Sci.*, 1985, **7**, 65–78.

5. Pupin, J. P., *Contrib. Mineral. Petrol.*, 1980, **313**, 207–220.
6. Larsen, L. H. and Poldervaart, A., *Min. Mag.*, 1957, **31**, 544–564.
7. Murthy, M. V. N. and Siddiquie, H. N., *J. Geol.*, 1964, **72**, 123–127.
8. Pidgeon, R. T. and Aftalion, M., *Geol. J.*, Spl. Paper, 1978, **4**, 183–220.
9. Srikantappa, C., Raith, M. and Spiering, B., *J. Geol. Soc. India*, 1985, **26**, 849–872.
10. Chacko, T., Ravindra Kumar, G. R., Meen, J. K. and Rogers, J. J. W., *Precambr. Res.*, 1994, **55**, 469–490.
11. Soman, K., Narayanaswamy, and Schmus, V. W. R., *J. Geol. Soc. India*, 1995, **45**, 127–136.
12. Narayanaswamy, and Soman, K., *J. Geol. Soc. India*, 1990, **35**, 46–64.
13. Narayanaswami, S. and Purna Lakshmi, *J. Geol. Soc. India*, 1967, **8**, 38–50.
14. Ravindra Kumar, G. R. and Chacko, T., *J. Geol. Soc. India*, 1986, **28**, 277–288.
15. Murthy, M. V. N., Siddiquie, H. N. and Viswanathan, T. V., *Zircon* (ed. Murthy, M. V. N.), Misc. Publn. 9, Geol. Surv. India, 1968, pp. 25–47.
16. Gastil, R. G., Delise, M. and Morgan, Jr., *Geol. Soc. Am. Bull.*, 1967, **78**, 870–906.
17. Poldervaart, A., *Zircon* (ed. Murthy, M. V. N.), Misc. Publn. 9, Geol. Surv. India, 1968, pp. 1–24.
18. Viswanathan, T. V. and Murthy, M. V. N., Charnockites, IGC Report of the 22nd Session, Part XIII, 1964, pp. 97–120.

ACKNOWLEDGEMENT. GRR thanks the Department of Science and Technology, New Delhi, for financial assistance.

Received 16 August 1995; revised accepted 6 November 1995

INDIAN INSTITUTE OF SCIENCE
BANGALORE 560 012

Post Doctoral Research Associateships in Life Sciences and Biotechnology
(A DBT-sponsord programme)

Applications are invited from candidates with Ph D in Biological, Chemical or Physical Sciences, or MD, for the award of temporary Research Associateship (maximum duration – 2 years) to carry out research in frontier areas of Biology in the Division of Biological Sciences at IISc under the DBT-sponsored postdoctoral programme. Candidates *who have submitted their theses* (awaiting award of degree) are also elibigle to apply, but they may be appointed as Research Associates (Provisional) till they obtain the degree. Minimum consolidated emoluments for Research Associates will be around Rs 4,000 p.m. This is an ongoing programme and the PDFs award will be operative from 1 April or 1 October of every year.

Candidates should apply on plain paper giving: (1) Biodata; (2) Summary of Ph D work including list of publications (reprints if any); (3) Two reference letters, from supervisor and persons familiar with their scientific work; and (4) An undertaking that they will continue with the training programme for its duration.

After initial screening, candidates will be called for an interview at IISc., Bangalore in the first/second week of March/September 1996. The PDF programme will start from 1st April/October 1996. Send applications to Prof. K. P. Gopinathan, Chairman, Department of Microbiology and Cell Biology, Indian Institute of Science, Bangalore 560 012 to reach him before **31 January 1996/15 August 1996.**

Prof. K. P. Gopinathan
Convenor
DBT-PDF Programme at IISc

Annual Review of Immunology 1995. William E Paul, ed. Annual Reviews Inc., 4139, El Camino Way, Palo Alto, California, 94303-0139, USA. Price: USA $ 48, elsewhere $ 53. Vol. 13. 729 pp.

The first 'personal essay' in the first volume of the *Annual Review of Immunology* in 1983 was written by Elvin Kabat, who watched over the progress of immune phenomena from physiology to biochemistry. The first actual review in that volume was written by Gustav Nossal, who writes, a dozen years later, the personal essay for this volume, and documents the progress of immune phenomena from the cellular to the molecular level. The difference between those two transitions is central to what has happened in immunology over the past decade, as a shift from the organism to its cells, and from biochemical to molecular events.

However, before examining this volume further—what are the *Annual Reviews* supposed to do, anyway? According to the editors, they keep those not directly working in the area 'up-to-date'. This is meaningful only if the 'updating' involves concepts, since few who do not work in an area are interested in its experimental facts. But, since the existence of the *Annual Review of Immunology* was justified in 1983 for its editors [who are still at its helm] by the sheer bulk of work in immunology, it is more likely that the reviews would be collections of technical facts rather than ground-breaking conceptual arguments, which is what happens, albeit with rare honourable exceptions. The reviews are compendia more than commentaries, and they seem to be content to be that, especially given, one supposes, the number of other 'review' periodicals in immunology that are pithier, shorter and 'laity-friendly'.

One result of this is that 'comprehensiveness' has become a special virtue for these volumes. This has led to two things. One, they simply tend to document the changing faces of experimental systems or 'phenomena' in immunology. This is aggravated by the fact that immunology is still a science that is frequently shifting its basic paradigms as the result of major experimental discoveries of 'phenomena', so that 'theories' seem more ephemeral to immunologists than they do, perhaps, to professionals in other, more 'established' sciences, although the allure of immunology also lies in that uncertainty. Two, this virtue of comprehensiveness has led the editors of the *Annual Reviews of Immunology* to declare [in the first volume itself] that appropriate scientists would be asked to write reviews on a broadly representative range of topics in immunology. This has meant that reviewed areas have not necessarily seen any major recent changes fundamental enough for the reviews to be anything but compendia of technical facts. All this makes the *Annual Reviews of Immunology*, and this volume no less than previous ones, boring to the casual reader from sister disciplines, and of utility mainly to workers looking for an easy survey of their field, with the pitfall that, since these are still reviews written by individuals, the 'comprehensiveness' may after all be idiosyncratic and less than all-embracing.

In judging the success of such a volume, therefore, all the poor reviewer can do for the lay reader is to survey what the volume does review, and, maybe, what possible lacunae there may be. But the latter is made difficult by the fact that the process of selecting topics and reviewers begins about two years before the publication of an Annual Reviews volume. Results of recent days that may, one thinks, merit a mention, such as the DM gene products, will therefore find space only in later years. This means that the wish of the editors of *Current Science*, that the book reviewer should essay a short review of the state of the art, would create an untenable situation where the book reviewer discusses issues that the volume itself does not mention.

After all this, what DOES this volume actually contain? Well, one can perhaps divide work that passes under 'immunology' into four broad areas: one, issues of interest uniquely to immunologists and for the moment to few other people, which usually means the generation and consequences of the clonally diverse recognition mechanisms that are the mainstay of the vertebrate immune system and are what classical immunologists mean by 'adaptive' immunity; two, issues that deal with a broader notion of 'defense' in eukaryotic organisms or what classical immunologists call 'innate' immunity; three, analyses that simply use the immune system to answer questions of general interest to all biologists such as in developmental regulation and signal transduction; and four, what everybody is interested in, how immunology explains, treats, prevents, or cures diseases. The present volume has a reasonable selection of subjects representing this entire range, in the proper tradition of 'political correctness'. In fact, it is instructive to look at how much of the present volume is dedicated to diseases—almost a half of it as opposed to barely a fifth in the first volume in 1983. This reflects both the fact that immunology has matured enough to be able to look at diseases without looking foolish, as well as the changing political climate for research funding everywhere, which has, in this era of post-modern conservatism, come to insist on immediate delivery of usable value for every penny spent on research.

Thus, adaptive immunity is represented by discussions about selection of the T cell repertoire, the effects of deletion of the MHC class II molecules in engineered mice, the basis of immune reactivity to so-called type 2 T-independent antigens, interleukin-12, and the relevance of CD19 in generating B cell responses. Innate immunity is represented by peptide antibiotics of mammalian origin, by the transcriptional regulation of complement genes, by the interactions of bacterial lipopolysaccharide with the immune system, and more dubiously, by a discussion of the effects of the neuro-endocrine axes on 'autoimmunity'. Issues of general interest in biochemistry and molecular biology discussed include the role of the *bcl-2* gene product in cell survival and death, signalling through cytokine receptors in haemopoietic cells, and potassium and calcium channels in lymphocytes. Reviews that purport to deal with diseases in the context of the immune system cover a very wide range, from situations where the disease is simply a tool for fundamental questions in adaptive immunity such as the essay on retroviral superantigens, through 'softer' issues such as the immune responses to *Leishmania major*, hepatitis B virus and HIV, via a discussion of the genetics of a so-called autoimmune disease animal model in the non-obese diabetic mouse and an attempt to link another disease 'model', reactive arthritis, to a more baffling disease, rheumatoid arthritis, finally to pieces that deal with therapeutic possibilities. These last possibilities reviewed are mostly conceptual right now, and even there many

of them are vague. They range from cytokine adjuvants for cancer immunotherapy to T cell therapy for human viral diseases and the possibilities in xeno-transplantation of organs, and they could be said to end in a notable review of the politics and mechanics of immunodiagnostic tests, in the example of HIV. This is an eclectic collection by any standards, and as would be expected, the quality varies, although few fall below the minimum required for coherence and 'comprehensiveness'. There is, in fact, a reassuring air of 'solidity' about the latter virtue.

There is, however, a sense of rote about many reviews despite being well-written, because they do not seem to stem naturally from exciting new breakthroughs in their area, but rather come as though an *a priori* decision had been made to include them for extraneous reasons. The sterling exception to that, and arguably the best review in the volume, is the review by Bevan and his colleagues on the selection of the T cell repertoire. It provides a terse and succinct background, sets up the theoretical paradox—that T cells need to recognize something during their development in order to be identified as useful and allowed to survive, but if they do recognize anything in the body at this time they are killed off in order to avoid the dangers of autoreactivity. Having stated the dimensions of the problem, the review then looks at the periphery of the problem, so to say, and examines the current status of ideas and data on, for example, the role of co-receptors in such selection and how these co-receptors are regulated during selection, before plunging into the genesis and testing of the various possible hypotheses that can explain the interactions between T cell receptors and their ligands that has made this so fermenting a field in recent years. The demonstration that peptides have a significant role to play in positive selection rather than only in negative selection of T cells, of landmark value enough by itself, is thus set in context and taken further to bold predictions of what may be happening *in vivo*—a peroration that fits well at the end of a rattling good read. Would that more reviews in this commodious compendium were as riveting.

SATYAJIT RATH

National Institute of Immunology,
Aruna Asaf Ali Marg,
New Delhi 110 067, India

Annual Reviews of Medicine 1995. Cecil H. Coggins ed. Annual Reviews Inc, 4139, El Camino Way, Palo Alto, California 94303-0139, USA. Vol. 46. Price: USA $ 47, Elsewhere $ 52. 549 pp.

Medical literature has become a big jungle these days, what with more than 30,000 biomedical journals pouring in data at a phenomenal pace. A novice in the field will find it difficult, if not impossible, to get at the rose wood and the teak wood in the jungle, instead he may get lost in the midst of the dead wood. Books like the medical annual will go a long way in helping people to know the useful research data and the milestone developments in their respective fields. This kind of book also could give the distilled wisdom in the field. 'Where,' said T. S. Elliot, 'is the wisdom we seem to have lost in knowledge?' and 'where,' he asked, 'is the knowledge we seem to have lost in information?' To convert the whole lot of information from the ordinary journals one needs to have a system whereby the juniors could get at the truth.

The editors of the medical annual have a very difficult task to pick and choose only a small fraction of the new knowledge generated during the year under review. The chapters in this book will, perforce, have to be patchy and choosy. However, this edition has been more than fair to the subjects under review. They have also selected the right people to do the reviews. People like Norman Kaplan, Erik Mogenson and Mark Pfeffer are all giants in their own fields, to mention only a few whose work I have direct experience of.

Molecular biology, new face of tuberculosis, brain imaging in schizophrenia, newer viruses, ventricular remodelling and organ transplantations are all the hot topics of the day and have all been judiciously reviewed. The index is a pleasure for the reader.

As in the previous issues of the medical annual which I have been following ever since 1964, the present one has kept up its high standards of the contents and the contributors. I had the pleasure of reviewing this good book two years ago for the *British Postgraduate Medical Journal* where I had expressed similar sentiments.

The book has been produced well, the print and the paper are also good. Overall this book should get very high marks and I have no hesitation to recommend it strongly. This must be compulsory reading for all the postgraduates and others in practice. I would have been happier if they had included a review of the newer problems in the field of tropical diseases, as some of them like malaria, have come back with a vengeance, and do not follow the old time pattern at all; as also the dangerous scenario of nosocomial infections.

The research into malaria and its molecular biology is more than just of practical importance. It has more to it than what meets the eye. An important current topic is the localization of genes implicated in the susceptibility to common chronic diseases. This has been greatly facilitated by the use of polymerase chain reaction. An interesting recent observation has been the apparent protection against autoimmune diseases like rheumatoid arthritis and systemic lupus erythematosus in areas of west Africa where malaria is endemic, in contrast to the high incidence of these in west African-origin Americans. If this is correct, it would be a good example of the old saying that 'autoimmunity is the price paid for eradicating infectious diseases'.

The book is highly recommended despite these remarks.

B. M. HEGDE

Kasturba Medical College,
Mangalore 575 001, India

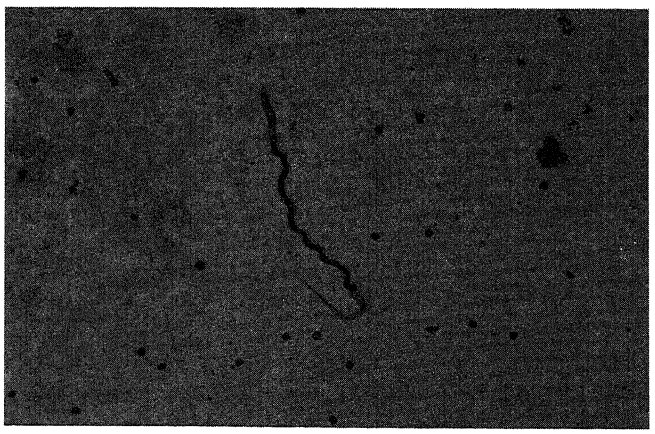
Tropical Diseases: Molecular Biology and Control Strategies. Sushil Kumar, A. K. Sen, G. P. Dutta and R. N. Sharma, eds. CSIR, New Delhi. 1994. 773 pp.

The prevalence of tropical diseases in India has some very impressive statistics: Forty million people are chronically infected with filariasis, 4 million with leprosy, 0.25 million with Leishmaniasis, 1.5 million tuberculosis cases, 1.0 million malaria cases and equally dreadful number of people with diarrhoeal infection in a year. Rough, probably conservative, estimates these are. As if these diseases are 'inseparable but terrible companion of man' in this country. The urgency to control these diseases cannot be over-

emphasized. Paradoxically, guidelines to control programmes are rather well known. These are chemotherapy, early diagnosis, development of vaccines, eradication of vectors and imparting health education to the community. But nothing could be easier said than done. Effective eradication of tropical diseases poses formidable challenges. There are 'friendly' vectors (mosquitoes, flies) to transmit some 'deadly' diseases like malaria, filaria, kala-azar, encephalitis, etc. Poor hygienic and socio-economic condition add to the misery and severity of the disease. (Parasitic diseases are unfortunately diseases of poor people.) In addition to a plethora of such problems, one is faced with pathogens getting resistant to drugs and vectors to insecticides. The organisms have developed sophisticated survival mechanisms to neutralize host's immune response, to act as 'smoke screen' so that host is unable to elicit protective immunity against the organism. (The process of parasitism is as old as life itself.) Answers to these problems can be found in understanding the fundamentals of host–parasite interactions. It is the belief of many researchers including the present reviewer that new strategies to control tropical diseases can be developed by the application and resources of biomedical sciences to pinpoint new drug targets, to develop diagnostics, to identify and clone protective antigens. It is indeed appropriate that CSIR, as a part of its Golden Jubilee celebration, brought together 250 scientists and research scholars working in tropical diseases for a national symposium held at Central Drug Research Institute, Lucknow (February, 1992), India. *Tropical Diseases—Molecular Biology and Control Strategies* is the title of the book, which consists of 93 papers presented at this symposium.

About half of the book is on filariasis and malaria. Discussion on filariasis begins with a paper on the natural history of infection followed by two papers on immunodiagnosis. The natural course of filariasis, which is a chronic disease, spanning decades needs to be established. The natural course of events starting from infection (bites of mosquitoes carrying infective larvae of the parasites) to the onset of clinical symptoms in humans will undoubtedly be complex, and needs to be studied. A good diagnostic should be able to 'detect' target population (be it asymptomatic microfilaraemic individuals or clinical filarial patients or even individuals 'exposed' to the infection). Early diagnosis of filariasis is a key area linked to control. Present efforts are directed in developing assay systems to detect microfilaraemic carriers. The two papers sum up the progress employing antibody and antigen assays. The latter system detects active infection (presence of parasite) in contrast to antibody assay which in principle cannot differentiate present from past infection. The last words on this difficult subject definitely have not been said. Two papers deal with protective mechanisms in filariasis using animal models. Extension of such studies to humans needs to be initiated. The drugs available for treatment and control of filaria are unsatisfactory and limited. The widely used Diethylcarbamazine lacks adult worm-killing activity, produces adverse reactions in humans, thereby limiting its use for mass therapy. The alternative, Ivermectin, is no better. There is thus a need to synthesize interesting molecules or to identify plant products that could be used as drugs in filariasis. Five chapters are about this subject. Vector control measures either for malaria, or filaria or kala-azar constitute a vital component in tropical diseases research. Four papers discuss this topic. The contents range from application of anti-vector agents in plant extracts, larvivorous fish (*Gambusia affinis*), insecticide impregnated bed nets, insecticides, to even physical removal of weeds (breeding sites) by community participation.

The ability of *Plasmodium falciparum*, the major parasite for human malaria, to overcome anti-malarial drugs is a major hindrance to the control of malaria both in individual and community levels. There are nine papers on the synthesis and mode of action of anti-malarial drugs. Efforts are made from CDRI, Lucknow, to synthesize better derivatives of Chinese drug Quinghasou (Artemesin) against cerebral malaria and new 8-aminoquinoline derivative (CDRI 80/53) against *P. vivax* infection. The biochemistry of malarial parasites and parasitized erythrocytes is a subject of extensive research. A set of eight papers addresses this issue. Enzymes of the parasites have been proposed to be the target of drug design, diagnostic and protective antigens. These assumptions rest on the premise that parasitic enzymes should differ substantially from host (mammalian) enzymes in structural and immunochemical properties. In most cases they are. Antibodies to lactate dehydrogenase of *P. falciparum* were demonstrated as serodiagnostic marker in malaria, polyamine oxidase as potential target for chemotherapy of protozoan diseases. The characterization of lactate dehydrogenase as a possible immunodiagnostic agent indicates an important finding. The production of reactive oxygen species by phagocytic cells is known to be a general defense mechanism against various pathogens. The role of free radicals and lipid peroxidation in malaria has been discussed in two papers. Development of vaccines for malaria with the usual steps of antigen identification, immunogenicity and adjuvant selection is fairly at advanced level compared to other parasitic diseases of man. Two papers report interesting work using synthetic peptides of RESA (ring infected erythrocyte surface antigen) and MSA (major



Microfilariae of human filarial parasite *Wuchereria bancrofti.*

merozoites surface antigens) of *P. falciparum*.

There are six papers on the visceral leishmaniasis or kala-azar. Immunodiagnosis will be highly useful in leishmania, diagnosis of which at present depends solely on biopsy materials. The procedure is painful and inconvenient. DOT-ELISA and direct agglutination tests have been recommended for immunodiagnosis and are discussed comprehensively in three papers. A paper describes the generation of monoclonal antibodies to *L. donovani* whose utility should have been explored. A set of thirteen papers discusses mycobacterial infections, tuberculosis and leprosy. Most of the papers are about the immunodiagnosis. A monoclonal antibody-based assay was described to be useful for diagnosing progressive form of leprosy. Paediatric population is the worst sufferer in any infection, tuberculosis is no exception. It is of interest to note that excretory secretory antigens of *M. tuberculosis* could be used as diagnostic antigen in detecting pulmonary and extrapulmonary tuberculosis in children. The diagnostic utility of ES antigen has parallel in the previous work of this group (Sevagram) in filariasis. A paper reported the lack of production of oxygen radicals from macrophages of leprosy patients and the consequent killing ability; which interestingly can be restored by delipidified cell component of *M. leprae*. Immunology of leprosy is often synonymous with poor delayed type hypersensitivity reaction and immune unresponsiveness. An easily cultivable mycobacterium, *M. habana*, has been reported to break this immunological tolerance. Development of vaccine has been a major focus in leprosy research. Immunotherapeutic potential of much known ICRC (Indian Cancer Research Centre) vaccine has been described in one paper. A single dose of vaccine can significantly upregulated *M. leprae*-specific immune responses in lepromatous leprosy patients.

Seven papers are about micro-organisms (*E. histolytica, Giardia lambia, Vibrio cholerae, E. coli*) that cause diarrhoeal disorders. Papers on the molecular biology of cholera and designing drugs based on molecular mechanism of diarrhoea are noteworthy. There are also two articles on 'ubiquitous' liposomes, as drug delivery vehicle specifically in malaria and in systemic fungal infections. As the role of liposomes in reducing the toxicity of drugs is now well known, its application at least in the later disease and kala-azar may well be accepted for humans in future.

A lone paper on Japanese encephalitis mention the generation of cytotoxic T lymphocytes and the nature of infected cell lysis. A paper on DNA probes describes the usefulness of the technology in detecting infections for *M. tuberculosis*. There are papers on the genetics of antibiotic resistance in pathogenic *A. cinetobacter*, a gram-negative bacteria capable of causing diverse infections and also on human oncogene expression.

The role of traditional medicines (Ayurveda) in the health care of a traditional country like India is very important. A paper describes the activity of plant extracts, selected on the basis of ethenobotanical information, against bacteria (*Klebsiella* sp., *E. coli*, *Shigella flexneri*) causing gastrointestinal and urinary tract infections. Clinical trial of 'Mahakhandaw churna' showed it to be very promising in intestinal amoebiasis compared to metronidazole.

There are some weaknesses which are inevitable perhaps in a book with such a broad range of topics with so many contributors. Some topics (disease) have been dealt reasonably comprehensively, some cursorily. I wonder if the editors could not have organized the papers under sections diseasewise for instance. The book is not free from printer's error either. Nevertheless the editors are to be congratulated for preparing the book which will be an useful addition to the libraries of all laboratories engaged in parasitic disease research. There is something for everyone here.

MANOJ K. DAS

Regional Medical Research Centre, Bhubaneswar 751 016, India

Edited and published by Prof. S. Ramaseshan and Prof. P. Balaram, Current Science Association, Bangalore 560 080.
Typeset by R D Typesetters, Bangalore 560 010. Printed by Lotus Printers, Bangalore (Ph: 3380167).

MEETINGS/SYMPOSIA/SEMINARS

Centenary Seminar on Applied X-Rays

Date: 18–19 December 1995
Place: Madurai

Topics include: Crystal growth and characterization; High-pressure X-ray diffraction; Crystal structures; Defect characterization; Imaging plate technique; Crystalline solid solutions; Other related topics.

Contact: Prof. S. K. Mohanlal
Convener, Centenary Seminar on Applied X-rays
Department of Physics, Madurai Kamaraj University
Madurai 625 021
Phone: 0452-85252
Fax: 0452-85239

DAE Homi Bhabha Chair for Distinguished Scientists/Professors including Engineers and Technologists

The Department of Atomic Energy has established the DAE Homi Bhabha Chair. There would be a total of not more than 8 such awardees of which 4 would be tenable at any of the research centres of the DAE and autonomous institutions aided by it and another 4 at Indian Universities or other academic and research institutions in the country.

Contact: Shri B. K. Saha
Joint Secretary (R&D)
Dept of Atomic Energy
Chhatrapati Sivaji Maharaj Marg
Bombay 400 039
Phone: 022-2022492 (O), 3631987 (R)
Fax: 022-2048476

International Workshop on Advanced Plastics and Rubber Composites (IWAPRC-96)

Date: 15–22 January 1996
Place: Kharagpur

In this workshop, a comprehensive discourse is envisaged on advanced polymeric composites comprising advanced plastics and elastomer matrices, newer fibres, manufacture of composites, mechanics of composites and structure-property, etc.

Contact: Prof. R. P. Singh/Prof. C. K. Das
Materials Science Centre
Indian Institute of Technology
Kharagpur 721 302, India
Phone: (03222) 2221-2224 extn.4792, 7792
Fax: 91-03222-2303
E-mail: rps.matsc.iitkgp.ernet.in

XIV All India Cell Biology Conference of the Indian Society of Cell Biology

Date: 23–25 February 1996
Place: Calcutta

Topics include: Cellular and molecular parasitology; Plant molecular biology and biotechnology; Human diseases: Diagnostics, mechanisms and therapeutic aspects; Animal reproduction and biotechnology; Structure-function of genes and chromosomes and cellular regulation.

Contact: Dr Dwijen Sarkar/Dr Pratap K. Das
Joint Secretaries
XIV All India Cell Biology Conference
Indian Institute of Chemical Biology
4 Raja S. C. Mullick Road, Calcutta 700 032
Tel: 033-473 6793/0492/3491/3493
Fax: 033-473 0350/5197
E-mail: iicb%sirnetc@sirnetd.ernet.in

National Conference on Recent Trends in Vibrational Spectroscopy

Date: 26-28 February 1996
Place: Meerut

Topics include: Vibrational spectroscopy of molecules; Electronic spectroscopy of molecules; Laser and laser-based spectroscopy; Vibrational analysis and molecular structure; Spectroscopic studies and molecular structure; Theoretical methods; Spectroscopy of metal-complexes; New techniques; Other related topics.

Contact: Dr V. K. Rastogi
KC-68/1, Old Kavi Nagar
Ghaziabad 201 002
Phone: 0575-840374
Fax: 011-2413388

Contact Programme for Excellence in Diagrammatics and Cartography in Earth Sciences

Date: 22–29 February 1996
Place: Pune

Topics include: Various aspects of diagrammatics in earth sciences and principles of representations; Computer aided graphics and demonstrations.

Contact: Dr Vivek S. Kale (Local Coordinator)
Department of Geology
University of Pune
Pune 411 007
Phone: 0212-356061 ext. 2115
Fax: 0212-479 306
E-mail: Vivek@unipune.ernet.in

XVIII Conference of the Indian Association for General Relativity and Gravitation (Dedicated to Prof. S. Chandrasekhar)

Date: 15–17 February 1996
Place: Madras

The programme will consist of plenary sessions, special session and sessions for contributed papers covering topics from classical gravity, cosmology, gravitational radiation, relativistic astrophysics, quantum gravity (including strings), etc. There will be a session on the late Professor Chandrasekhar.

Contact: Dr G. Date
Institute of Mathematical Sciences
C.I.T. Campus, Tharamani
Madras 600 113
Tel: 044-235 1856
Fax: 044-235 0586
E-mail: sham@imsc.ernet.in

First International Conference on Molecular Association

Date: 18–20 March 1996
Place: Aligarh

Molecular association (including organic conductors, semiconductors and super conductors) with special section on biological, environmental and biomedical applications and chemical education is being organized at Aligarh.

Contact: Dr Pushkin M. Qureshi
Conference Director, ICMA 96
Laboratory of Molecular Association
Department of Chemistry, AMU
Aligarh 202 002
Tel (Day) 0571-400515
T. Fax: 009(571) 400076
Telex: 564-230-AMU-IN

Sheet Metal Working '96

Date: 22–23 October 1996
Place: Hanover, Germany

The Conference Theme: Process-and quality assurance in sheet metal working and tool and die making.

Contact: Verein Deutscher Ingenieure VDI
VDI-Gesellschaft Produktionstechnik (ADB)
Postfach 10 11 39
D 40002 Dusseldorf
Germany
Fax: +49211 6214 575

Physical and Biological Changes across the major Geological Boundaries

Date: 10–12 September 1996
Place: Lucknow

Topics proposed: Precambrian-Cambrian Transition; Permo-Triassic boundary; Cretaceous-Tertiary boundary; Pleistocene-Holocene boundary.

Contact: Dr P. K. Maithy
Convener
Birbal Sahni Institute of Palaeobotany
53, University Road
Lucknow 226 007

All India Institute of Medical Sciences

Ansari Nagar, New Delhi 110 029

Training in Electron Microscopy for Scientific Investigators for a period of 21 days w.e.f. 19.2.96 to 12.3.96

A training programme for a duration of three weeks will be held in the Electron Microscope Facility at the All India Institute of Medical Sciences, Ansari Nagar, New Delhi 110 029.

Applications are invited on plain paper stating the following:

(1) Name, (2) Age, (3) Postal Address, (4) Educational Qualification(s), (5) Name of the University/ Institution graduated from, (6) Positions held with current designation, (7) Whether Permanent/Temporary, (8) A brief write-up on the nature of work and why training in Electron Microscopy is considered necessary, and (9) Letter of recommendation from the present employer. Applications giving the above information should reach Dr A. K. Susheela, Officer-in-Charge, Electron Microscope Facility, All India Institute of Medical Sciences, New Delhi 110 029, **within 15 days of publication this Advertisement.**

A fee of Rs. 3000/- will be charged for the selected candidates. The mode of payment will be indicated in the Letter of Selection.

BIOINFORMATICS PROJECT

SCHOOL OF BIOTECHNOLOGY

MADURAI KAMARAJ UNIVERSITY

Applications are invited on plain paper for the following post in the Bioinformatics Centre, School of Biotechnology, Madurai Kamaraj University, Madurai.

Position: Junior Information Scientist (Technical)
Scale of pay: Rs 2200–4000, with usual allowances

Qualification and Experience

Candidates with a post-graduate degree in Computer Science or Physics or Electronics or Biophysics required. Basic qualification could be relaxed in the case of candidate with exceptional experience.

Knowledge of computer hardware and networking essential. Experience in VAX-VMS, Unix-operating systems and language programming preferred. Familiarity with biology-related software will be an advantage. The candidate will be required to manage the computer systems and maintain and extend the thick ethernet network connectivity VMS, DOS and Unix operating system. The job entails user service-related activities.

The appointment is for the scheme period ending 31 March 1997.

Five copies of the application along with the bio-data should be addressed to the Registrar and marked to the Scientist-in-Charge, Bioinformatics Centre, School of Biotechnology, Madurai Kamaraj University, Madurai 625 021 within fifteen days from the date of this advertisement.

REGISTRAR

We'll help you get there

If you're selling science books/journals or laboratory products, or wish to recruit scientists in your establishment, then

you're sure to benefit by advertising here.

Because *Current Science* reaches nearly every university, scientific institution and industrial R&D unit in India. What's more, it's read by hundreds of individual subscribers – students, doctoral scientists and professionals in virtually every field of scientific activity in India.

Current Science has the largest circulation among scientific journals in India.

And to give you more impact, we're now bigger, and better.

Write now, or send your copy to

CURRENT SCIENCE

C.V. Raman Avenue, P.B. No. 8001,
Bangalore 560 080.

NEW Display Advertisement Rates*

		Tariff (rupees)					
		Inside pages		Inside cover pages		Back cover page	
No. of insertions	Size	B&W	Colour	B&W	Colour	B&W	Colour
1	Full page	5,000	10,000	7,000	12,000	10,000	15,000
	Half page	3,000	5,000	–	–	–	–
6	Full page	25,000	50,000	35,000	60,000	50,000	75,000
	Half page	15,000	25,000	–	–	–	–
12	Full page	50,000	1,00,000	70,000	1,20,000	1,00,000	1,50,000
	Half page	30,000	50,000	–	–	–	–
24	Full page	90,000	1,80,000	–	–	–	–
	Half page	50,000	90,000	–	–	–	–

SEND PAYMENT BY BANK DRAFT WITH ORDER

*Effective from 1 January 1995

CURRENT SCIENCE

A fortnightly journal of research

Editors: S. Ramaseshan and P. Balaram

Editorial Board

S. Arunachalam, Central Electrochemical Research Institute, Karaikudi 623 006
Asis Datta, School of Life Sciences, Jawaharlal Nehru University, New Delhi 110 067
D. Balasubramanian, Centre for Cellular and Molecular Biology, Hyderabad 500 007
J. Chandrasekhar, Department of Organic Chemistry, Indian Institute of Science, Bangalore 560 012
R. Gadagkar, Centre for Ecological Sciences, Indian Institute of Science, Bangalore 560 012
K. N. Ganeshaiah, Department of Plant Genetics and Breeding, University of Agricultural Sciences, Bangalore 560 065
R. V. Hosur, Tata Institute of Fundamental Research, Bombay 400 005
N. V. Joshi, Centre for Ecological Sciences, Indian Institute of Science, Bangalore 560 012
Jyotirmoy Das, Indian Institute of Chemical Biology, Calcutta 700 032
C. C. Kartha, Division of Cellular and Molecular Cardiology, Sree Chitra Tirunal Institute for Medical Sciences and Technology, Thiruvananthapuram 695 011
S. S. Krishnamurthy, Inorganic and Physical Chemistry Department, Indian Institute of Science, Bangalore 560 012
S. C. Lakhotia, Department of Zoology, Banaras Hindu University, Varanasi 221 005
K. Muralidhar, Department of Zoology, Delhi University, Delhi 110 007
R. Narasimhan, CMC Limited, Bangalore 560 001
K. Neelakantan, ANURAG (Advanced Numerical Research and Analysis Group), Hyderabad 500 258
R. Nityananda, Raman Research Institute, Bangalore 560 080
T. Padmanabhan, Inter-University Centre for Astronomy and Astrophysics, Pune 411 007
B. R. Pai, Propulsion Division, National Aerospace Laboratories, Bangalore 560 017
T. J. Pandian, School of Biological Sciences, Madurai Kamaraj University, Madurai 625 021
Y. S. Rajan, Department of Science and Technology, New Delhi 110 016
M. Ramakrishnan, Geological Survey of India, Calcutta 700 016
A. V. Ramani, T. T. K. Pharma Ltd., Bangalore 560 025
S. Ramaswamy, Department of Physics, Indian Institute of Science, Bangalore 560 012
A. K. Raychaudhuri, Department of Physics, Indian Institute of Science, Bangalore 560 012
S. R. Shetye, Physical Oceanography Division, National Institute of Oceanography, Dona Paula, Goa 403 004
V. Siddhartha, 51 Bharati Nagar, New Delhi 110 003
V. Srinivas, School of Mathematics, Tata Institute of Fundamental Research, Bombay 400 005
G. Srinivasan, Raman Research Institute, Bangalore 560 080
R. Uma Shaanker, Department of Crop Physiology, University of Agricultural Sciences, Bangalore 560 065
K. S. Valdiya, Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore 560 064
G. Venkataraman, Department of Physics, Sri Sathya Sai Institute of Higher Learning, Prasanthi Nilayam 515 134
M. Vidyasagar, Centre for Artificial Intelligence and Robotics, Bangalore 560 001
K. VijayRaghavan, TIFR Centre, Indian Institute of Science Campus, Bangalore 560 012

Current Science, founded in 1932, is published by the Current Science Association in collaboration with the Indian Academy of Sciences. The journal is also intended as a medium for communication and discussion of important issues that concern science and scientific activity. All articles published in *Current Science*, especially editorials, opinions and commentaries, letters and book reviews, are deemed to reflect the individual views of the authors and not the official points of view, either of the Current Science Association or of the Indian Academy of Sciences.

Working Committee, Current Science Association

P. Balaram, Indian Institute of Science, Bangalore
R. Chidambaram, Department of Atomic Energy, Bombay
V. K. Gaur, CSIR Centre for Mathematical Modelling and Computer Simulation, Bangalore
K. Kasturirangan, ISRO Satellite Centre, Bangalore
P. K. Kaw, Institute for Plasma Research, Gandhinagar
N. Kumar, Raman Research Institute, Bangalore (*Vice-President*)
N. V. Madhusudana, Raman Research Institute, Bangalore (*Secretary*)
A. Mani, Field Research Unit, Indian Institute of Tropical Meteorology, Bangalore (*Past-President*)
R. A. Mashelkar, Council of Scientific and Industrial Research, New Delhi
R. Narasimha, Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore
G. Padmanaban, Indian Institute of Science, Bangalore (*Vice-President*)
R. Ramachandran, The Institute of Mathematical Sciences, Madras
P. Rama Rao, Defence Research and Development Organization, New Delhi
S. Ramaseshan, Raman Research Institute, Bangalore (*Treasurer*)
M. S. Valiathan, Manipal Academy of Higher Education, Manipal (*President*)

Editorial Office

Editorial staff

Chandrika Ramesh, G. Madhavan, T. D. Mahabaleswara, C. S. Ravi Kumar, M. S. Venugopal

Circulation and Accounts

Peter Jayaraj, Ranjini Mohan, B. Sethumani, Shanthi Bhaskar, B. K. Shivaramiah

Annual subscription 1995

	Institutional	Personal
India	Rs 350	Rs 100
All other countries (AIR MAIL)	US $175	US $50

Single copies other than special issues: Rs 35/US $15

Editorial Office: CURRENT SCIENCE
C. V. Raman Avenue
P.B. 8001
Bangalore 560 080, India
Telephone: 91-80-3342310
Telex 845-2178 ACAD IN
Fax: 91-80-3346094

Copyright © 1995, Indian Academy of Sciences

INFORMATION FOR CONTRIBUTORS

GENERAL

All manuscripts should be addressed to the Editor, *Current Science*, P. B. No. 8001, C. V. Raman Avenue, Bangalore 560 080. Submission of an article will be held to imply that it has not been previously published and is not under consideration for publication elsewhere; and further, that if accepted, it will not be published elsewhere. *Three copies of contributions of all categories* are required, with a letter of transmittal giving (i) names and complete addresses of the authors and (ii) title of the contribution and the category in which it is submitted (see below).

Current Science is a multidisciplinary journal and therefore research and review papers of general significance that are written clearly and well organized will be given preference. All papers will be first assessed by a Reviewing Editor. Papers found unsuitable in terms of the overall requirements of the journal will be returned to the authors. The others will be sent for detailed review. *Both solicited and unsolicited material will be reviewed.* Authors of these papers will be notified of acceptance, rejection, or need for revision of the paper. Returned papers cannot be resubmitted.

Illustrations and other material to be reproduced from other publications must be properly credited; it is the authors' responsibility to obtain permission for reproduction (copies of letters of permission should be sent).

CATEGORIES OF MANUSCRIPT

General articles (not exceeding 5000 words) discuss current trends in research in a field that will be of interest to readers outside the field; interdisciplinary topics: science policy and administration; or some aspect of the application of science and technology to human needs or the impact of science and technology on society/ecosystems/life. They should include a summary not exceeding 100 words, introductory paragraph(s), brief subheads at appropriate places to point to what follows, illustrations that will help a general reader, and references.

Review articles (not exceeding 5000 words) are expected to survey and discuss recent developments in a field. They should be well focused and organized, and avoid a general, 'textbook' style.

Research articles (not exceeding 4000 words) should report research results of fairly major significance. They should include an abstract not exceeding 100 words, introductory paragraph(s), and brief subheads.

Research communications (not exceeding 2000 words) should contain important findings that are novel and of fairly broad interest. They should include a brief abstract and an introductory paragraph. Text should not be broken up under subheads.

Correspondence includes letters, not exceeding 500 words, that are of general interest to scientists. All letters cannot be published.

Scientific correspondence contains technical comments, including those on articles or communications published in *Current Science* within the previous six months. Letters may be reviewed and edited.

Research news articles are intended to inform nonspecialists about recently published advances or important findings discussed at a meeting. Authors should also send a copy of the paper(s) on which the article is based. Meeting reports should avoid merely listing brief accounts of topics discussed, and must convey to readers the significance of an important advance.

Research accounts articles are intended to be personalized reviews of research from the authors' own laboratory, based on a body of published work. The articles must provide appropriate background to the area in a concise introduction, which should also serve to place the author's work in proper perspective. Articles will normally not exceed 8 to 10 printed pages.

Opinion articles present views on issues related to science and scientific activity. **Commentary** articles should contain expository notes on issues related to science and scientific activity.

Book reviews. Unsolicited reviews will also be considered. Reviews that merely 'list' brief descriptions of the contents cannot be published. Reviews should have 'context' and convey some information about the subject of the book.

Historical commentary and notes inform readers about interesting aspects of personalities or institutions of science or about watershed events in the history/development of science; most are expected to relate to India. Illustrations are welcome. Brief items will also be considered.

MANUSCRIPT PREPARATION

Manuscripts should be typed double-spaced on one side of white bond paper (21 × 28 cm). The pages should be numbered consecutively, starting with the title page and through the text, reference list, tables and figure legends. The **title** should be brief, specific and amenable to indexing. Not more than five **keywords** should be indicated separately; these should be chosen carefully and must not be phrases of several words. **Summary** and **abstract** should not have more than 100 words and should convey the main point of the paper, outline the results and conclusions, and explain the significance of the results.

Text. All papers should have a brief introduction. The text should be intelligible to readers in different disciplines and technical terms should be defined. Tables and figures should be referred to in numerical order. All **symbols** and **abbreviations** must be defined, and used only when absolutely necessary. Superscripts and subscripts and ambiguous characters should be clearly indicated. **Units of measure** should be metric or, preferably, SI. Methods should, as far as possible, be described briefly in appropriate table and figure legends.

Figures. In the case of line drawings one set of originals (without any lettering) is sufficient, with two copies containing lettering. In the case of photographs good prints are required with each copy of the manuscript; photocopies are not acceptable. Line drawings should be roughly twice the final printed size. The correct orientation should be indicated if not clear.

Photomicrographs and other photographs that require it must have a scale bar, which should be defined clearly in the legend. Primary data should be submitted as far as possible (e.g. actual photographs of electrophoretic gels rather than idealized diagrams).

References should be numbered in the order in which they appear, first through the text and then through table and figure legends. The following are examples of ways of writing references.

1. Mukundan, T. and Kishore, K., *Curr. Sci.*, 1991, **60**, 355–362.
2. Constantine, G., in *Biology of Bats* (ed. Wimsatt, W. A.), Academic Press, New York, 1970, vol. 1, pp. 319–322.

Acknowledgements should be brief. Footnotes are not allowed except to identify the corresponding author if not the first.

Cover photographs. Good photographs (colour or black and white) that pertain to a submitted paper will be considered for use on the cover. Good *prints* and a legend should be submitted with the manuscript. In the case of a colour picture, a transparency will be required for printing if accepted.

PROOFS AND PUBLICATION

Two sets of galley proofs are sent to the corresponding author. A reprint order form accompanies the proofs.

ıunded: 1932

ISSN: 0011–3891

CURRENT SCIENCE

olume 69 Number 12

25 December 1995

955 **In this issue**

CORRESPONDENCE

957 **A crisis in engineering education**
S. C. Dutta Roy

957 **Vitamin C: A vital nutrient**
C. S. Kavitha

958 **NSU, Academy document and science education: Reflections and an operation model**
A. Bhaskara Rao and V. K. Garg

NEWS

959 **Nobel leptons**
K. V. L. Sarma

963 **Rice Biotechnology – Report on the Third International Rice Genetics Symposium October 16–20, 1995, Manila, Philippines**
C. R. Bhatia

RESEARCH NEWS

964 **Designer molecules: A chromogenic sensor of creatinine**
Uday Maitra

965 **Pathogenesis: Generating new paths in understanding bacterial virulence**
Kakoli Mukhopadhyay

COMMENTARY

968 **Are herbaria redundant?**
R. R. Rao

SCIENTIFIC CORRESPONDENCE

969 **The academy fellowship problem**
J. V. Narlikar

971 **Supplementary comments on 'endemic, rare and threatened flowering plants of South India'**
M. P. Nayar; R. J. Ranjit Daniels, N. Anil Kumar and M. Jayanthi

971 **A fast growing cell line of *Taxus baccata* L. (Himalayan yew) as a potential source of taxol precursor**
Sumita Jha and Timir B. Jha

GENERAL ARTICLES

974 **Solar photovoltaic energy products from semiconductor physics to advanced energy technology**
Ashok Parthasarathi, D. Madhavan and V. K. Kaul

986 **Mushrooms: Beauty, diversity, relevance**
C. V. Subramanian

RESEARCH ARTICLE

998 **Role of fluorescence microscopy in the assessment of Indian Gondwana coals**
Bhagwan D. Singh

RESEARCH COMMUNICATIONS

1005 **Reduced uptake of (^{3}H) thymidine by PHA-responsive human peripheral blood lymphocytes as a function of storage**
Awadhesh N. Jha

1008 **3.55 Ga old zircon from Singhbhum–Orissa Iron Ore Craton, eastern India**
J. N. Goswami, S. Mishra, M. Wiedenbeck, S. L. Ray and A. K. Saha

1012 **Significance of the terminal Cretaceous calcareous nannofossil marker *Micula prinsii* at the Cretaceous–Tertiary boundary in the Um Sohryngkew section, Meghalaya, India**
Rahul Garg and K. P. Jain

1017 **Origin and significance of high grade phosphorite in a sediment core from the continental slope off Goa, India**
V. Purnachandra Rao, Ch. M. Rao, M. Thamban, R. Natarajan and B. R. Rao

1023 **Palynofloral analysis of sections from Bharathi and Kundara clay mines of Kerala Basin: Palaeoecological and tectonic perspective**
K. P. N. Kumaran, K. Soman, C. V. Kamble and Ancy Joseph

1028 **Bioremediation of toxic metal ions from polluted lake waters and industrial effluents by fungal biosorbent**
Md. Naseem Akthar and P. Maruthi Mohan

BOOK REVIEWS

1031 **Deep continental structure of India: A review,** *reviewed by* Ram S. Sharma

1031 **Rice Tungro,** *reviewed by* V. Muniyappa

1032 **Bacterial pathogenesis of plants and animals: Molecular and cellular mechanisms,** *reviewed by* J. S. Virdi

ERRATUM

1033 **Pioneers of ozone study bag chemistry Nobel**
S. Parthiban [*Curr. Sci.*, 1995, **69**, 793–797]



COVER. *Amanita muscaria*, the fly-agaric. See page 986.

Indexed in CURRENT CONTENTS/GEOBASE/CHEMICAL ABSTRACTS

In this issue

Solar photovoltaics

In India as soon as one hears the phrase 'solar energy' the mind slips into a well-oiled groove. Visions of 'appropriate technology' arise. Unwieldy structures of wood and galvanized iron covered with aluminum foil in which blackened vessels are kept to boil rice seem to dot the open countryside. Unfortunately, no low cost method of utilizing solar energy has succeeded commercially. The best way to use solar energy is to convert it to electricity. While cost of photovoltaics as such is a major deterrent worldwide, the lack of the basic technology itself is of major concern in India.

India now seems poised to take its place in the sun. The Central Electronics Limited (CEL) under the Department of Scientific and Industrial Research, Ministry of Science and Technology has made giant strides in the indigenous development of photovoltaic technology. CEL recently won a prestigious international tender against stiff competition from developed countries.

Ashok Parthasarathy has always been a champion of indigenous technology. Under his guidance, both the National Research Development Corporation and the CEL have done a lot to commercialize indigenous technologies for exploitation not only in India but abroad as well. Parthasarathy *et al.* describe in their article (**page 974**) how they traversed the thorny path from the Physics Lab to the state-of-the-art photo voltaics industry.

Amorphous silicon cells cost less but have a low efficiency. Polycrystalline silicon cells are much more efficient but at a much higher cost. Thin film polycrystalline silicon cells show great promise of combining high efficiency with a relatively low cost.

India has a vast potential for tapping solar energy. The efforts of CEL should further us considerably towards achieving this goal.

A. V. Ramani

Mushrooms: Science and myth

Mushrooms, a complex and advanced group of fungi, have been the centre of folklore, legends and magic from time immemorial. They have been a source of misconception and superstition in many a society, probably because of their association with putrescence and decay and because of their often unpleasant colour and fetid smell. As a group, they excel other plants in their sheer diversity of form, smell, colour and function. They have some of the most enviable names such as earth ball, the blusher, the false blusher, lawyer's wig, death cap etc. Botanically the mushrooms that we see soon after a rain, are the fruiting bodies of the fungi. While most of the mushrooms are saprophytic, a few form beneficial association with other plants and fewer still are parasitic.

Amanita muscaria, commonly known as the fly-agaric is one of the most colourful mushrooms and is known for its hallucinogenic property. This mushroom has entered into the folklore and ritual of all major religions. Our own 'soma rasa' is believed to be a cocoction derived from this fungus. In fact the Rigveda contains hymns in praise of this concoction and is said to accurately depict the morphological characteristics of the fly-agaric. However, not all mushrooms impart such intoxication. Many contain nearly fatal toxins, which have been used time and again in history to avenge defeat. Infact on this John Gerard, as early as 1597, warned as mildly as one could 'a few of them are good to be eaten and most of them do suffocate and strangle the eater!' In his article (**page 986**), C. V. Subramanian reviews the fascinating story of these mushrooms blending the science and the myths. A world of opportunity to ponder over and work in these exciting systems is thus unravelled.

R. Uma Shaanker

Current Science on the Internet

The contents pages of *Current Science* have been available on the World-Wide-Web for the last several months. They can be accessed from the home page of the Centre for Ecological Sciences, Indian Institute of Science, Bangalore, using

http://ces.iisc.ernet.in (or http://144.16.65.194).

To go directly to the home page of *Current Science*,

http://ces.iisc.ernet.in/Current Science

can be used. Contents pages for all the issues of *Current Science*, beginning with the 10 July 1995 (up to 10 November 1995) can be found at the site. Also shown are images of the cover pages of 25 August 1995 and 10 September 1995 issues.

There have been over 150 visits to the *Current Science* home page over the last couple of months (excluding those from the many departments within the Indian Institute of Science), for viewing both the contents pages and the cover pages. This includes accesses from Canada, Germany, Sweden, Netherlands, Japan as well as several educational, commercial, governmental and non-governmental organizations in the USA and other parts of the world.

If at all there are any installations which have only ftp access to the internet (and are unable to access the World-Wide-Web), the contents pages can also be downloaded from the anonymous ftp site of the Centre for Ecological Sciences (ces.iisc.ernet.in).

CURRENT SCIENCE

Volume 69 Number 12 25 December 1995

CORRESPONDENCE

A crisis in engineering education

There exists an acute shortage of high-quality teachers in engineering institutions, particularly in the high technology areas. Young talents are not attracted to teaching because of the long drawn process of acquiring the minimum required qualification (a Ph D for IITs, IISc, etc./M Tech for others) and then starting at the same or perhaps a lower salary than what he/she would have got had he/she chosen to join industry immediately after the first degree (B Tech). True, there are exceptions, but they are so insignificant in number that they do not count. Also, even if a young talent joins the teaching profession, the career prospects are rather limited, because the salary of the highest available position in the profession, viz. that of a full professor compares rather poorly with that of a comparable person in industry. This gap has widened, rather than shortened, over the years, particularly in areas like electronics and computers. The situation is such that most of the engineering institutions have to make do with whatever is available in the market, and this is far from being a happy situation. Moreover, adequate manpower, even of questionable competence, is not available so that there exists, always, a significant number of vacant positions in any institution. Obviously such manpower cannot produce high-quality graduates even if the entrants to the courses are of high quality. The process, if allowed to continue, will create, or has already created, a vicious circle. For the advancement of technology, and for technological progress of the country, effective and concrete solutions to the problem must be found and implemented – the sooner, the better!

What constitutes an effective solution? To retain high-quality retiring teachers for a few more years? This has been, and is being tried in many places. But then this is, at best, a temporary measure and cannot be a permanent solution. The obvious solution is to make the teaching profession a very prestigious one – in monetary as well as other terms so that the most talented young people are attracted to this profession, rather than to civil service or industry. Let the level of achievement demanded for the lowest position be high, but let it be justified not only on academic grounds but also in terms of monetary benefits. This will need a substantial change in the attitude of the Government and other funding agencies towards education to the extent that they are willing to allocate to education, a budget which compares not too unfavourably with that of some other important sector (like defence, for example?).

Shortage of good teachers in engineering education has reached a crisis point and this fact deserves serious concern and attention of all concerned.

S. C. DUTTA ROY

Department of Electrical Engineering, Indian Institute of Technology, Delhi 110 016, India

Vitamin C: A vital nutrient

The review article published by Indu B. Chattejee and coworkers (*Curr. Sci.*, 1995, **69**, 747–751) on 'Vitamin C: A potential saviour against free radical-induced oxidative damage' is really a beautiful one, picturizing the story of an essential nutrient which has varied roles in human nutrition. The article has effectively pointed out the importance of vitamin C and how it functions in order to prevent the various degenerative diseases and aging. Vitamin C is one of the most neglected nutrients in our dietaries. The essential and rich sources of vitamin C are green leafy vegetables and fruits, which are often neglected. In a country like ours where cereal-based diet is the staple one, availability of this vitamin can only be through the inclusion of green leafy vegetables (GLV) and fruits, which has been drastically reduced compared to the earlier days due to the blind emulation of westernized diet patterns, where high-calorie diets play an important role, with relatively low vitamin C content. This is really a disappointing trend as we have rich edible plant diversity, with good vitamin C content. The recent tendency towards animal foods, which are low in vitamin C content, should be checked and ample thrust should be given to the inclusion of GLV and fruits, atleast a minimum of 50–75 g daily, in our food pattern. The best way would be the raw form as far as possible because vitamin C is heat labile and easily destroyed during prolonged heating, to reap the full potential of this vital nutrient.

Indeed the most important reasons for developing many degenerative diseases are improper diet pattern and sedentary life style. But inclusion of the essential nutrients can improve the situation to some extent.

C. S. KAVITHA

Krishi Vigyan Kendra, National Research Centre for Spices, Peruvannamuzhi, P.O., Calicut 673 528, India

NSU, Academy document and science education: Reflections and an operation model

India, as any other developing country, is passing through a transformation in all spheres of activity. Education, research and the society have much to suffer or get benefited by it. As NRIs, visiting regularly India and maintaining contacts with Indian Universities, we are aware that *suggestions and proposals need to be considered on their feasibility*. Otherwise, we continue debating eternally.

Perhaps our tone of discussion should change and a few specific questions addressed.

1) How competitive do we prepare our youth to face problems and seek solutions?

2) What are the expectations of the industry we are attending to?

3) How good is the output of our research to the society?

4) How much of time do we spend to prepare ourselves to update our own knowledge to pass it on with confidence to the students?

5) How often had we time for self-evaluation and self-criticism?

6) How well can we stimulate youth towards confidence in what they learn and in what they will face in the profession?

7) How much of time do we spend to guide and inspire a student to fetch exciting ideas, and encourage his ideas, without copying them and discouraging them?

8) How can we convince ourselves that we are part of a social upbuilding and that our contribution in our activity is important for the country?

9) How do we solve these problems?

There is no harm in having the input from people and Academies. Only it is necessary to see if they could be feasible *within the existing structure and without prejudicing the ongoing activities in the nation*. Expecting everything from the Government is unrealistic. Thus we also like to give some solutions to one of the ideas, with the hope that our country could forge ahead with its traditional high capacities of talents and creativity.

NSU is feasible

1) Indian science needs reshuffling to activate it further and promote environment for the generation of intelligent ideas and discoveries.

2) This is possible with the feeding in of some new brains and new ideas of action, proposed by an NRI. The spirit behind the idea is to stimulate a supposedly static system in operation and control over Indian science.

3) NRIs are spread all over the world. They love India as others in India do. They also want to contribute to Indian efforts to attain better goals in development. As we see, this is the spirit behind this NSU proposal.

4) Any increase in Universities or Institutes should not interfere in the priorities of the government towards the needs of the large mass of citizens. The fact that all the product of these new Centres of Excellence would revert its genius towards the benefit of the nation, is not doubted, though it is a question of time.

5) It is absolutely feasible to create a group of NSUs spread over India within a short period through the direct involvement of a NRI Foundation.

6) The NRI Foundation will be established in the USA. It will be fed and controlled by the NRIs. Its main activity is to promote the installation, operation and maintenance of the NSUs, one after another in India. Thus NSUs will be Foundation Universities, nonprofit organizations, but operate as Private Institutions. The surplus in the annual budgets will be reverted to NRI Foundation for its activities.

Modus operandi

We are sure that by making proposals within the parameters of their feasibility, all our discussions will be very constructive. Thus we believe that NSUs are absolutely possible and would be very successful. When they adopt a NGO structure, access to the NSUs will be through merit alone.

It is easy to make proposals and offer suggestions. When convinced, it is not difficult to take a lead in trying to make them a reality. (Sometimes theory and experiment are both correct but do not agree with each other; sometimes a wrong theory agrees with experiment. One must therefore be careful not to jump to conclusions[1].) This depends on the desire to succeed, and with a personal commitment. The authors on proposing this NRIF operational system are convinced of its success and thus announce their decision to be among the very first to become Supporting Members of the Foundations.

1. Aharoni, A., *Physics Today*, June 1995, **48**, 33.

A. BHASKARA RAO*
VIJAYENDRA K. GARG†

**Department of Mineral Resources,*
†Department of Physics,
University of Brasilia
70910-900 Brasilia, Brazil

Nobel leptons

This is the year of the leptons. Of the six leptons of today's particle physics, Nobel Prizes were awarded earlier for discovering two of them: the electron and the muon-neutrino. This year it is the turn of two more leptons, the electron-neutrino and the tau lepton, to acquire 'Nobelity'.

Frederick Reines (University of California, Irvine), aged 77, reported the first direct evidence for the existence of the 'neutrino' in 1956. He performed the experiments jointly with Clyde L. Cowan, Jr., who died in 1974. The neutrino is a tiny neutral particle hypothesized by Wolfgang Pauli, Jr. way back in 1930, and was thought to be unobservable. Twenty five years later and with the invention of the nuclear reactor it was possible to verify the neutrino hypothesis.

Martin Perl (Stanford University, Stanford), aged 68, assisted by a 35-member team, discovered the 'tau lepton' in 1975. This particle carries a unit of electric charge and weighs approximately twice as much as the proton. It is characterized by a short life of only a fraction of a pico-second and has the distiction of being the heaviest and the shortest-lived lepton. It is interesting that while Reines's neutrino results were long hoped for, Perl's tau discovery came as a complete surprise – that too at a time when people were busy establishing the charm flavour around the same mass. The tau lepton heralded the existence of the 'third generation' of fundamental constituents of matter.

A brief history of the neutrino

> The neutrino is the smallest bit of material reality ever conceived of by man; the largest is the universe. To attempt to understand something of one in terms of the other is to attempt to span the dimension in which lie all manifestations of natural law. Yet even now, despite our shadowy knowledge of these limits, problems arise to try the imagination in such an attempt.
>
> F. Reines and C. L. Cowan, Jr.[1].

Neutrino 'was born' in a letter written by Pauli. This historic letter dated 4 December 1930 from Zürich, started with the unusual greeting 'Dear Radioactive Ladies and Gentlemen', and was read out at the Tübingen meeting. Pauli mentioned that he had hit upon the neutrino as a 'desperate remedy' to save, among other things, the principle of energy conservation in beta decay. The letter referred to the hypothetical particle as 'the neutron' (the nuclear constituent ~~neu~~tron was discovered later by James Chadwick in 1932). In Italian the word 'neutrone' means something like 'large neutral one'. It was Enrico Fermi who named the Pauli particle as the 'neutrino', meaning the little neutral one, and that name stuck on.

In 1931 Pauli visited Rome and discussed his proposal of the particle with Fermi who received the idea with 'a very positive attitude'. Official announcement of the neutrino by Pauli was made at the Solvay conference in October 1933, as a particle which possessed no electric charge but carried the missing energy and momentum and escaped the detecting equipment. The famous beta decay theory of Fermi appeared in 1934.

It is interesting that even the great Pauli did not fully recognize the implications of the neutrino, particularly in regard to its penetrating power. His own account of this early period, written sometime in 1957, is extremely fascinating and is now available in English translation[2]. Initially he thought he had done a 'frightful thing' as the neutrino was expected to have penetrating power similar to, or about 10 times larger than, a gamma ray. However in 1934 Hans Bethe and Rudolf Peierls argued that the neutrino had to be even more elusive as its interaction mean free path had to be astronomical in magnitude (note 1).

It was Bruno Pontecorvo who first suggested that the process of inverse beta decay could be used as a way to establishing the neutrino. In 1946 he proposed the celebrated Chlorine–Argon radiochemical method and suggested using the source of reactor neutrinos[2] for the experiment. The Cl–Ar method was taken up almost as a life-long project by the solar-neutrino pioneer Raymond Davis, Jr. (note 2).

Detection of the reactor neutrino

The first neutrino reaction observed (note 3) was the inverse nuclear beta reaction

$$\bar{\nu}_e + p \rightarrow e^+ + n, \qquad (1)$$

driven by the antineutrinos from the nuclear reactor. This reaction is essentially the reverse of the neutron decay $n \rightarrow p + e^- + \bar{\nu}_e$ from which it is obtainable by transposing the e^- to the left hand side as a e^+ and reversing the reaction arrow. The energy of the incident antineutrino could be deduced from the positron momentum:

$$E_{\bar{\nu}_e} = \frac{E_e + \Delta + [(\Delta^2 - m_e^2)/2M_p]}{1 - [(E_e - p_e \cos\theta)/M_p]} \qquad (2)$$

$$\simeq E_e + \Delta \qquad (3)$$

$$\geq 1.804 \text{ MeV}, \qquad (4)$$

where E_e is the positron energy, p_e is the magnitude of its 3-momentum, θ is the positron emission angle with respect to the antineutrino direction, and $\Delta = M_n - M_p = 1.293$ MeV is the neutron–proton mass difference; the second step ignores neutron recoil terms of order (E_e/M_p) and the last one gives the reaction threshold.

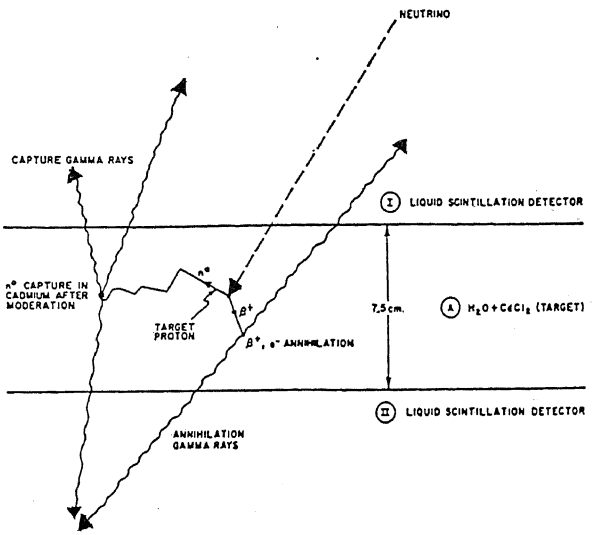
For MeV range of antineutrino energies, the positron angular distribution is nearly isotropic ($\langle\cos\theta\rangle = -0.04$) and the effective cross section may be taken to be (note 4)

$$\sigma = \frac{4G_F^2}{\pi} p_e E_e, \qquad (5)$$

where G_F is the Fermi coupling constant; for $E_{\bar{\nu}_e} = 2.3$ MeV the cross section is

$$\sigma \simeq 6\times10^{-44} \text{ cm}^2. \qquad (6)$$

The idea underlying the Nobel Prize-winning experiment was to look for a pair of scintillator pulses, the first (prompt) pulse due to positron annihilation and the second (delayed) one due to capture of the moderated neutron. The experiment[3] was performed around 1955–56. The projectiles were the reactor neutrinos from the Savannah River Plant located in South Carolina State, USA, and targets were the protons in a solution of water mixed with cadmium chloride (Cd is a good absorber of thermal neutrons). The experimental apparatus could be viewed as a multiple-sandwich with 2 layers of cheese (target material) arranged between 3 pieces of



Schematic diagram of neutrino detector.

Figure 1. Schematic diagram of the neutrino detector of Reines–Cowan experiment[4].

bread: the cheese layers were the two target tanks containing a solution of water and $CdCl_2$, while the bread pieces were the three liquid scintillator detectors.

An event meant the detection of two prompt coincidences (see Figure 1): the first one was between the two photons (each having 0.511 MeV energy) of the positron annihilation, and the second prompt coincidence was due to the capture of neutron by cadmium giving a few photons (with total energy of 9 MeV). The second pulse occurred after several micro-seconds of the positron flash, the time it took for the neutron to be thermalized in the target water. The experiment involved, among other things, measuring the energies of the pulses, their time-delays, dependence of event rate with neutrino flux which varied with the reactor power output, etc.

The results were published in the 5-author paper C. L. Cowan, Jr. *et al.*[3] entitled 'Detection of the Free Neutrino: A Confirmation'. This paper, unusual for the lack of diagrams of the detector or the customary figures or tables, is regarded as the discovery paper (note 5). Photographs of the detector assembly and associated equipment can be seen in ref. 4. The observed signal varied with the reactor-power. It consisted of an average rate of 2.88 ± 0.22 counts/hour, consistent with a value $\sim 6 \times 10^{-44}$ cm^2 for the inverse beta-reaction cross section. The signal to reactor-unrelated background ratio was 3 to 1.

It is remarkable that in the same year Reines and Cowan[1] also gave an upper limit on the neutrino magnetic moment,

$$\mu_{\bar{\nu}_e} < 10^{-9} \mu_B,$$

where μ_B is the Bohr magneton. This limit was deduced from the extent of non-observation of scintillator pulses along the path of the reactor neutrino, and is about the same as the present upper limit which is extracted by using different considerations (for a survey of the recent results in neutrino physics, see e.g., ref. 5).

Even today, we do not know very much about the electron-neutrino (and much less about other neutrinos). As for its mass, data on the end-point of the beta electron spectra only show that it does not exceed a few eV; a recent high statistics experiment[6] using molecular tritium in gaseous form, gave the upper limit

$$m(\bar{\nu}_e) < 4.35 \text{ eV}, \quad (7)$$

at 95% confidence level.

Experimental data with ν_es are not extensive because it is difficult to obtain ν_e beams. At the pi-meson factories, low energy ν_e get produced in 3-body decays and along with muon-neutrinos (ν_μ and $\bar{\nu}_\mu$). Sun is no doubt a good source of ν_e but solar neutrino fluxes are not well understood and constitute one of the challenging problems of current research. Only recently the first experiment using a man-made source of pure ν_e was performed. It made use of an intense source of reactor-produced ^{51}Cr (which emits ν_e by electron capture) to calibrate the GALLEX solar neutrino detector[7].

Discovery of the tau lepton

The tau lepton broke on the scene unexpectedly. While the results of Reines needed the construction of power reactors, the discovery of tau lepton needed high energy electron–positron colliders. Tau lepton is the third kind of charged lepton that exists in Nature, the other two being the electron and the muon. (The Greek letter τ is the first in the word triton, meaning third). Its birth can be traced to the 1975 paper entitled, 'Evidence for anomalous lepton production in $e^+ - e^-$ annihilation' by Perl *et al.*[8].

The experiment was performed at the electron–positron collider called SPEAR (Stanford Positron Electron Accelerator Ring). At this facility, beams of e^- and e^+ were accelerated simultaneously in opposite directions in a ring and made to intersect. As the total center-of-mass energy (the sum of beam energies)

$$E_{cm} = E_{e^-} + E_{e^+} \quad (8)$$

$$= 2E_{e^-}, \quad (9)$$

was tunable in the range 3–8 GeV, a pair of charged particles each having a rest mass of about 2 GeV could easily be produced at SPEAR.

A large cylindrical detector placed in a magnetic field surrounded the collision area. Electrons were identified by the electromagnetic shower counters, and muons by their ability to penetrate large amounts of iron and other materials making up a total of 1.7 absorption lengths for pions. Particles emitted at polar angles between 50° and 130° and at all azimuthal angles were recorded.

Anomalous events. The 'anomalous' events reported by Perl and collaborators corresponded to the following reactions which had a very distinctive signature,

$$e^- + e^+ \rightarrow e^- + \mu^+ + \text{i.p.}, \quad (10)$$

$$\rightarrow e^+ + \mu^- + \text{i.p.}, \quad (11)$$

where 'i.p.' denotes invisible particles which left no trace in the detector. The ingenuity of the experimenters consisted in establishing that the oppositely-charged $e\mu$ pair arose from the separate decays of two new particles which were oppositely-charged and short-lived. In an effort to reduce the background due to the copious production of e^+e^- and $\mu^+\mu^-$ pairs, the events were chosen to be 'acoplanar' so that production of more than two particles was ensured. For this the μ was required to make an angle more than 20° to the plane that contained the final e and the incident beam.

The very first report of Perl *et al.*[8] had a total of 64 anomalous events at the SPEAR range of energies. For example, at the energy $E_{cm} = 4.8$ GeV there were 24 events (13 $e^+\mu^-$ and 11 $e^-\mu^+$) with an estimated background consisting of about 5 events (arising from possible misidentification of hadrons as leptons, decays of known hadrons into leptons, etc). The events were found to be 'noncollinear', meaning that the angle between the e and μ momenta was more than 90° and the two particles were emitted in opposite hemispheres with respect to the beam.

Threshold behaviour. The occurrence of anomalous events as a function of E_{cm} exhibited an increase around 4 GeV, as shown in Figure 2. This indicated the existence of a threshold for the production of anomalous events. Moreover, in producing a pair of point-particles $\tau^+\tau^-$ by one-photon exchange

$$e^+ + e^- \rightarrow \gamma_{virtual} \rightarrow \tau^+ + \tau^-, \qquad (12)$$

quantum electrodynamics tells us that the total cross section $\sigma_{\tau\tau}$ should depend on the final velocity β as follows:

$$\sigma_{\tau\tau} = \frac{(3-\beta^2)}{2}\beta\sigma_0 \quad (\text{for spin } \tfrac{1}{2}), \qquad (13)$$

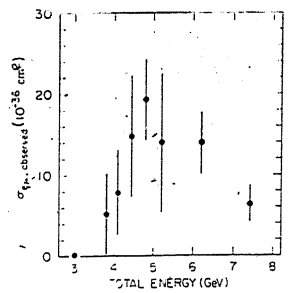


Figure 2. The observed cross section for the anomalous e–μ events[8].

$$= \frac{\beta^3}{4}\sigma_0 \quad (\text{for spin } 0), \qquad (14)$$

where

$$\beta = \sqrt{1 - \frac{4M_\tau^2}{E_{cm}^2}}; \quad \sigma_0 \equiv \frac{4\pi\alpha^2}{3E_{cm}^2}; \quad \alpha \simeq \frac{1}{137}. \qquad (15)$$

In the boson case, the β^3 dependence arises from the p-wave production which is a result of parity conservation in electromagnetic interactions. Subsequent experimental observations were consistent with a linear rise of the cross section with velocity and thus for the $\tau^\pm$ the assignment of spin $\frac{1}{2}$ was preferred over spin 0 and other possibilities.

Final interpretation of the anomalous events followed soon[9]: electron–positron annihilation gives rise to a pair of tau leptons

$$e^+ + e^- \rightarrow \tau^- + \tau^+ \qquad (16)$$

which decay immediately into lighter leptons. The 'tau–lepton number' conservation is respected by assuming the emission of an associated neutrino called the tau-neutrino (ν_τ) in the τ^- decay, and a tau-antineutrino ($\bar{\nu}_\tau$) in the τ^+ decay. Thus to explain the anomalous event with, say, $e^-\mu^+$ we should appeal to the decays

$$\tau^- \rightarrow e^- + \bar{\nu}_e + \nu_\tau, \qquad (17)$$

$$\tau^+ \rightarrow \mu^+ + \nu_\mu + \bar{\nu}_\tau, \qquad (18)$$

wherein the two neutrinos and the two antineutrinos (belonging to all three neutrino generations) constitute the invisible particles.

Data gathered over the years have shown that the shape of the electron (and also muon) energy spectrum in tau decay is in good agreement[10] with the standard V–A theory. Also all the observations support the view that the e–μ universality is extendible to e–μ–τ universality. Thus the original enigma regarding the existence of muon is deepened. It now becomes the so-called 'generation puzzle': among the ultimate constituents of matter, why do members of one generation behave exactly in the same way, except for mass, as the corresponding members of another generation?

It may be interesting to recall that in 1974 the spectacular discovery of J/ψ (3.097 GeV) immediately brought forth in its wake an intense frenzy of activity relating to the charm flavour. People were busy studying the spectroscopy of hidden-charm states and open-charm states and their decays. Consequently the feeling prevalent at the time was that Perl's anomalous events were some obscure manifestation of charm decays and the situation would be clearing up soon. However such feelings were firmly dispelled by 1977 when the Double Arm Spectrometer (DASP) group[11] working at DESY, Germany, reported seeing the anomalous $e\mu$ events around ψ' (3.686 GeV) manifestly below the open-charm threshold (2×1.87 GeV).

The present best value for the tau lepton mass comes from the measurements at the Beijing Electron Positron Collider (BEPC)[12], and is given by

$$m_\tau = 1776.96^{+0.18+0.25}_{-0.21-0.17}\ \text{MeV}, \qquad (19)$$

where the first error is statistical and the second systematic. The relatively large mass of the tau implies a large phase space for decay, and this makes a variety of final states to be accessible for the tau decay. Some of the decay states contain 5 charged particles (pions or kaons) besides the missing tau-neutrino and possibly one or two neutral pions. The availability of several competing channels for decay (e.g., the relative decay probability for $\tau^- \rightarrow \pi^- + \nu_\tau$ is $(11.7 \pm 0.4)\%$), makes the tau a very shortlived particle. Its meanlife is presently known to 1% accuracy[13],

$$\tau_\tau = (2.956 \pm 0.031) \times 10^{-13}\ \text{s}. \qquad (20)$$

Finally, in regard to the mass of the tau-neutrino, the information available is very limited: the mass limit obtained from a recent study[14] of the tau-decay events each containing 5 charged pions is, $m(\nu_\tau) < 24$ MeV at 95% confidence level.

Summary

The announcement of this year's Nobel award in Physics is in recognition of two landmark experiments in elementary particle physics. One provided the first confirmation of the neutrino, as envisaged by Pauli more than two decades earlier, as a tiny neutral particle emitted in beta decay. The experiment made an ingenious use of time-correlations of scintillator pulses to ascertain the occurrence of the inverse scattering reaction.

Table 1. Lepton generations

Generation	Lepton	Discoverer(s) (year)	Nobel Prize winner(s) (year)
1	electron, e	J. J. Thomson (1897)	J. J. Thomson (1906)
	electron-neutrino ν_e	C. L. Cowan *et al.*[3] (1956)	F. J. Reines (1995)
2	muon μ	J. C. Street, E. C. Stevenson; C. D. Anderson, S. H. Neddermeyer – (1936)	
	muon-neutrino ν_μ	G. Danby *et al.*[16] (1962)	M. Schwartz, L. M. Lederman, J. Steinberger (1988)
3	tau τ	M. L. Perl *et al.*[8] (1975)	M. L. Perl (1995)
	tau-neutrino ν_τ	?	?

The other experimental investigation which shared this year's honour is the discovery of the heavy charged lepton τ. This was a serendipitous discovery perhaps like that of the muon. The 'anomalous' events with the oppositely-charged e and μ were interpreted as resulting from the independent decays of the shortlived pair $\tau^+\tau^-$ which was created in the reaction $e^+ + e^- \rightarrow \tau^+ + \tau^-$ The τ is a member of the third (and perhaps the last) generation of ultimate constituents of matter to which the top and bottom quarks belong (see e.g. ref. 15).

A summary of the discoveries made in the world of leptons is given in Table 1. We see that the third generation has started getting Nobel prizes. It is amusing that the charged-leptons crop up with a 39-year gap and may be the 4th one would show up in the year 2114. For the present, the available experimental information implies that there are no charged leptons which are heavier than the tau and lighter than 45 GeV.

Finally it should be emphasized that the third kind of neutrino ν_τ still needs to be identified experimentally (this is shown by question marks). To this end one ought to demonstrate that τs are produced directly in collisions of the ν_τ with nuclear targets. However ν_τ beams are hard to obtain mainly due to the short intrinsic lifetime of the τ.

1. Reines, F. and Cowan, Jr., C. L., *Nature*, 1956, **178**, 446–449.
2. Pauli, W., in *Neutrino Physics* (ed. Winter, K), Cambridge University Press, 1991, pp. 1–25.
3. Cowan, Jr., C. L., Reines, F., Harrison, F. B., Kruse, H. W. and McGuire, A. D., *Science*, 1956, **124**, 103–104.
4. Reines, F. and Cowan, Jr., C. L., *Phys. Today*, 1957, **10**, 12–18.
5. Sarma, K. V. L., *Int. J. Mod. Phys.*, 1995, **A10**, 767–783.
6. Belesev, A. I. *et al.*, *Phys. Lett.*, 1995, **B350**, 263–272.
7. Anselmann, P. *et al.*, *Phys. Lett.*, 1995, **B342**, 440–450.
8. Perl, M. L. *et al.*, *Phys. Rev. Lett.*, 1975, **35**, 1489–1492.
9. Perl, M. L. *et al.*, *Phys. Lett.*, 1976, **B63**, 466–470.
10. Perl, M. L. *et al.*, *Phys. Lett.*, 1977, **B70**, 487–490.
11. Brandelik, R. *et al.*, *Phys. Lett.*, 1978, **B73**, 109–114.
12. Bai, J. Z. *et al.*, SLAC-PUB-95-6930, BIHEP-EPI-95-004, July 1995, (submitted to *Phys. Rev. D*).
13. Montanet, L. *et al.*, *Phys. Rev.*, 1994, **D50**, 1173–1827.
14. Buskulic, D. *et al.* (ALEPH Collaboration), *Phys. Lett.*, 1995, **B349**, 585–596.
15. Sarma, K. V. L., *Curr. Sci.*, 1995, **68**, 605–611.
16. Danby, G. *et al.*, *Phys. Rev. Lett.*, 1962, **9**, 36–44.

Notes

1. For the case of reactor antineutrinos going through water, taking the cross section to be $\sigma(\bar{\nu}_e + p \rightarrow e^+ + n) \sim 6 \times 10^{-44}$ cm^2 and noting that there are two free proton targets per water molecule, the interaction mean free path will be $\lambda = (1/n\sigma) \simeq 2.5 \times 10^{20}$ cm $\simeq 260$ light years!
2. It is worth recalling that the Cl–Ar experiment of Davis performed in 1956 at the Savannah River Plant reactor was a 'failure'. The result was not clearly inconsistent with background. The null result $\sigma(\text{reactor } \nu + {}^{37}\text{Cl} \rightarrow e^- + {}^{37}\text{Ar}) < 0.9 \times 10^{-45}$ cm^2 was interpreted to demonstrate that the reactor neutrino is not ν_e but $\bar{\nu}_e$. Since $\bar{\nu}_e$'s are emitted in the beta decays of neutrons and neutron-rich nuclei, a fission reactor is an abundant source of $\bar{\nu}_e$.
3. Here as is customary the word 'neutrino' is being used in its generic sense although, strictly speaking, what Reines and his group detected were the signals from the electron–antineutrino $\bar{\nu}_e$.
4. At MeV energies this has an approximate quadratic dependence on the incident neutrino energy E_ν; at higher energies (above GeV) it becomes independent of E_ν due to the rapid decrease of the vector and axial-vector form factors. However the total cross section increases linearly with E_ν, as a consequence of the point nature of the target quark. At still higher energies (where the W mass effects become important) the total cross section is expected to increase logarithmically, as is indicated in the recent HERA experiments $ep \rightarrow \nu_e + \ldots$.
5. The earlier 1953 experiment of Reines and Cowan, *Phys. Rev.*, 1953, **90**, 492; **92**, 830) performed at the Hanford reactor in Washington, did not give a signal well above the background that was unrelated to the reactor. The number of delayed coincidences had a large uncertainty, 0.41 ± 0.20 per hour.

K. V. L. Sarma, Tata Institute of Fundamental Research, Bombay.

Rice Biotechnology – Report on the Third International Rice Genetics Symposium October 16–20, 1995, Manila, Philippines

The Third International Rice Genetics Symposium was held at Manila, Philippines from October 16–20. The symposium was organized by the Rice Genetics Cooperative, International Rice Research Institute (IRRI), Los Banos and supported by the Rockefeller Foundation. IRRI has been organizing the rice genetics symposia every five years since 1985. In the last 12 years the Rockefeller Foundation has been intensively supporting rice biotechnology research with an investment of 65 million dollars. These funds have supported rice biotechnology research in the Asian countries as well as encouraged the advanced plant molecular biology laboratories in Europe and USA to take up research on rice. Further close linkages between the laboratories in the developing and developed countries have been established. As a result, rice plants are now grown in glass houses, with the backdrop of the Alps, in Switzerland and rice research is carried out at Cornell and Wisconsin Universities in USA and at John Innes Centre, Norwich and the University of Birmingham (UK). Japan has invested heavily in the Rice Genome Research Programme. China has a modest Rice Genome Programme besides a large number of laboratories engaged in rice biotechnology. This investment in rice has resulted in the blooming of rice biotechnology. This was clearly visible in the symposium with over 500 participants and over 300 oral and poster presentations. A marked shift towards molecular biology approaches was evident in comparison to the previous two symposia held in 1985 and 1990. In fact, at the concluding session, the participants were reminded by Ralph Riley that molecular biology only provided additional, new tools to be used by the plant breeders to facilitate their work of developing more productive, stress-tolerant cultivars with better enduse qualities.

Papers were presented in the areas of rice genome characterization, genetic transformation of rice, molecular tagging and map based cloning, gene isolation characterization and expression, quantitative trait loci (QTL) mapping, molecular cytogenetics, molecular analysis of cytoplasmic male sterility, genetic diversity in pathogen populations, genetics of morphological, physiological, disease resistance and quantitative traits. Gene mapping and transformation emerged as the two dominant areas of current world-wide interest in rice biotechnology.

Among the cereals, rice happens to have the smallest genome size and this has made it an ideal material for molecular biology studies and biotechnology. Restriction fragment length polymorphism (RLFP) technology was applied initially in the rice genome mapping but now polymerase chain reaction (PCR) based markers are being increasingly used. It was reported that 1904 markers consisting of cDNA clones as well as RAPD markers have been mapped covering a distance of 1590 cM. Integration of cytological, conventional and RFLP linkage maps was reported resulting from correct orientations of the maps with respect to each other.

Synteny (similar arrangement of genes along the chromosomes) between rice and wheat chromosomes was reported earlier. This has been further confirmed now and extended to barley, maize, sorghum, sugarcane, foxtail millet, pearl millet, oats and ryegrass. This makes the cloned and mapped genes of rice extremely useful for isolating and mapping genes in other cereals and should provide a quantum leap for molecular biology of other cereals.

Conversion of cloned RFLP markers of the Cornell University – IRRI genetic map to sequence-tagged sites (STS) was reported. The use of STS facilitates the breeding programmes by way of speed, convenience, reliability and relatively low cost as the PCR methods are used for analysis. Marker-assisted breeding would greatly facilitate selection for insect and pathogen resistance and in pyramiding of resistance genes.

Early results on association between molecular markers and quantitative trait loci (QTL) which determine the components of yield and ultimately the yield were reported. Such markers can assist in evaluation of the germplasm, selection of parents-to-be used in hybridization programme, prediction of hybrid vigour and in making selections.

A large number of reports were presented on the transfer and expression of alien genes into rice providing resistance to insects, viral, bacterial, and fungal diseases and nematodes. Reliable *Agrobacterium* mediated and gene gun methods have been developed. For insect resistance synthetic, codon optimized Cry 1A(b), Cry III A, from *Bacillus thuringiensis*, soybean trypsin inhibitor, cowpea trypsin inhibitor, potato proteinase inhibitor (pin 2), snow drop (*Gallanthus nivalis*) agglutinin (GNA) and ribosome in activating protein; for virus resistance, rice yellow mottle virus coat protein, hoja blanca virus (RHBV) and tungro virus, chitnase and cercosporin B gene for fungal diseases and herbicide resistance *bar* gene; and the gene for glycene betaine synthesis to provide salinity stress tolerance have been transferred to rice. Stability and expression of some of the transferred genes have been checked in the progenies upto the fifth generation. Time from the transformation event to obtaining the transformed plants and harvest of seeds from such plants have been considerably shortened. Approximately 11 weeks to obtain plantlets and less than a year for harvesting seeds were reported. Transformed plants are under glass house evaluation and are expected to move out into the field shortly. Insertion of single copy of the gene at one site, or multiple copies at one site as well as single or more than one copies at multiple sites have been observed.

In one presentation, transfer of rice clones encoding proteins from MADS-domain were transferred and expressed in tobacco where they altered apical dominance resulting in early flowering and bushy plants. Expression of Os-MADS3 altered floral organ development and homeotic development of petals to stamen. These results establish that the regulatory factors for flowering in dicots and monocots may be similar. Such homeotic genes have been extensively investigated in *Arabidopsis*.

The other significant reports described efforts made towards cloning of blast resistance genes by chromosome walking, presence of transposable elements in rice as well as in the pathogen

for blast disease, encoding of a receptor-like protein kinase by disease resistance gene, molecular analysis of anthocyanin pigmentation pathway, manipulation of starch metabolism during seed development, characterization of mitochondrial genome, silencing of β-glucuronidase gene in transgenic rice.

It has been projected that by the year 2025, the global demand for rice would increase by 70%, requiring production of 810 million tons of unmilled rice compared to the present production of 480 million tons. This will have to be harvested from diminishing cultivated area and availability of water due to competing demands on these resources. It was repeatedly emphasized that to meet this demand the rice yields will have to increase from the present mean of 4.5 tons ha^{-1} to 8.0 tons ha^{-1} under irrigation and from 1.9 tons ha^{-1} to 3.6 tons ha^{-1} in the rainfed areas. This will have to be achieved by increasing the yield ceiling, bridging the gap between the potential and the realized yield and sustaining the present levels. The major role of biotechnology towards this goal would be in the development of transgenic plants with resistance to biotic and abiotic stresses. Therefore there is an urgent need to evolve procedures for field releases of transgenic plants and approval of grain harvested from such plants as safe food similar to those produced using conventional genetic tools. The marker-assisted selection and tissue culture methods would facilitate breeding programmes and reduce the breeding cycle time.

C. R. Bhatia, Department of Biotechnology, New Delhi.

RESEARCH NEWS

Designer molecules: A chromogenic sensor of creatinine

Uday Maitra

Specific recognition of molecules by one another via shape complementarity and specific, weak non-covalent interaction like hydrogen-bonds, van der Waal's interaction, etc. is central to biology. Recognition of a chosen guest by specifically designed host molecules is currently a major area of chemical research.

Molecular and ionic sensors have always been in great demand for the easy detection/assay of a variety of molecules and ions[1]. Such sensors find many applications in medical diagnostics and in detecting environmental pollutants. A variety of chemical sensors, including biosensors, can be designed[2]. However, sensors based on synthetic molecular receptors are becoming more attractive since such molecules are usually much more stable compared to biomolecules. Of sensors, chromogenic sensors are the most valuable ones since the presence of the 'guest' is signalled by a change in colour. A number of chromogenic sensors have recently been designed and synthesized based on the principles of molecular recognition[3].

In a recent communication to *Science* (1995, **269**, 671), Bell and coworkers have designed a receptor which detects creatinine efficiently. Creatinine is one of the end products of nitrogen metabolism, and in healthy human beings it is excreted in the urine. The transport of creatinine from the blood to urine is mediated by the kidneys, hence the blood level of creatinine is a key indicator of renal function. Present methods available for the assay of creatinine are tedious and expensive. Bell's work has produced a simple but sensitive assay of creatinine, in which the formation of a supramolecular complex is signalled by a colour change.

The H_2N–C=N moiety of creatinine has a D-D-A array* for forming hydrogen bonds. The designed receptor however, in its neutral form, has a D-A-A surface for forming an H-bond. However, a zwitterionic structure is also possible for this receptor, in which the phenolic hydrogen is transferred to the nitrogen atom at the other end of the molecule, thereby converting the D-A-A surface to an A-A-D surface which is now complementary to the creatinine surface. In other words, of the two tautomeric forms only the zwitterionic form (the 'dye'-like form) should bind creatinine. In this example, that is what happens – the receptor exists predominantly in its neutral form, say in dichloromethane solution, which has a yellow colour. However, when this solution is shaken with an aqueous solution of creatinine, the receptor extracts creatinine into the organic solution, forms the complex, and in that process undergoes a change from the neutral to the zwitterionic form. This causes a colour change from yellow to brownish orange (the complex shows an extra absorption band at λ_{max} = 444 nm). This is schematically shown in

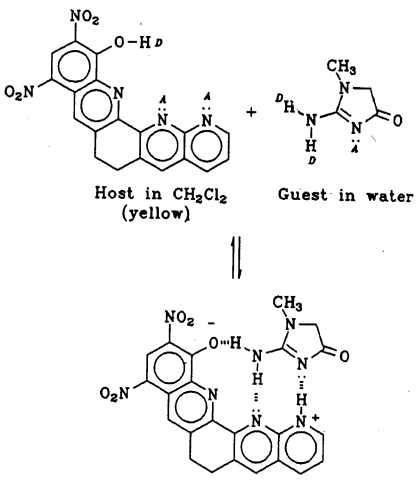


Figure 1

*D and A stand for donor and acceptor groups for forming a hydrogen bond, respectively.

Figure 1. The authors have also carried out X-ray structural analysis of the receptor, creatinine, and the complex to demonstrate some of the finer details of the binding.

The authors have shown that this synthetic receptor is selective for creatinine, since no colour change was observed with creatine, urea, uric acid, proline or histidine. The estimated dissociation constant of the complex (0.5 μM) is also sufficiently low so that blood serum creatinine (typically 40–130 μM) can be sensed by it. Alkali metal ions were found to give some chromogenic response with the receptor, but those were significantly weaker. The prospects for the use of synthetic receptors in medical diagnostics are bright.

1. For a recent review: Fabbrizzi, L. and Poggi, A., *Chem. Soc. Rev.*, 1995, 197.
2. *Chemical Sensors*, (ed. Edmonds, T. E.), Blackie, London, 1988.
3. Sandanayake, K. R. A. S., Imazu, S., James, T. D., Mikami, M., Shinkai, S., *Chem. Lett.*, 1995, 139; Zhao, S. and Luong, J. H. T., *J. Chem. Soc., Chem., Commun.*, 1995, 663; Rojas, M. T. and Kaifer, A. E., *J. Am. Chem. Soc.*, 1995, **117**, 5883; Sandanayake, K. R. A. S. and Shinkai, S., *Angew. Chem. Int. Ed. Engl.*, 1994, **33**, 2207; Fabbrizzi, L., Licchelli, M., Pallavicini, P., Perotti, A. and Sacchi, D., *Angew. Chem. Int. Ed. Engl.*, 1994, **33**, 1975; Vance, D. H. and Czarnik, A. W., *J. Am. Chem. Soc.*, 1994, **116**, 9397; Chawla, H. M. and Srinivas, K., *Chem. Commun.*, 1994, 2593; Hamasaki, K., Ikeda, H., Nakamura, A., Ueno, A., Toda, F., Suzuki, I. and Osa, T., *J. Am. Chem. Soc.*, 1993, **115**, 5035; Teixidor, F., Flores, M. A., Escriche, L., Vinas, C. and Casabo, J., *Chem. Commun.*, 1994, 963; Marsella, M. J. and Swager, T. M., *J. Am. Chem. Soc.*, 1993, **115**, 12214.

Uday Maitra is in the Department of Organic Chemistry, Indian Institute of Science, Bangalore 560 012, India

Pathogenesis: Generating new paths in understanding bacterial virulence

Kakoli Mukhopadhyay

The survival strategy of bacterial pathogens requires establishment in a suitable niche within another living organism, outwitting it by avoiding its formidable immune system, acquiring nutrients from it and multiplying, thus causing disease in the host. The host–pathogen interaction is a battlefield where the infecting organism encounters a range of adverse environmental conditions. These, along with other selective pressures, appear to have directed the evolution of specialized regulatory systems controlling the expression of virulence factors in the pathogen. To understand mechanisms by which the bacterial parasites circumvent or avoid the immune system of the host and cause disease, it is necessary to identify and characterize the bacterial genes and gene products responsible for the virulence at each stage of the development and progression of the disease. Traditionally, the strategies for identification of virulence factors included the three following approaches:

(1) Brute force screens, that pinpoint the virulence genes on the basis of mutational loss of a virulence phenotype[1,2].
(2) Cloning screens, that pick up novel virulence genes depending on their ability to confer a virulence-associated phenotype in a heterologous bacterium[3].
(3) Regulatory screens, that identify potential virulence genes by their coordinate expression with other known virulence genes under defined laboratory conditions[4,5].

All these strategies, carried out *in vitro* were, however, very often limited by their inability to reproduce accurately the complex and changing environments encountered by the pathogens in their hosts. Moreover, there could also be certain factors which would be induced only under specific environments within a host and could never be expressed, or were expressed in extremely low and negligible amounts, *in vitro*. Therefore it remained quite possible for one to miss out these genes if only the traditional methods were followed.

To overcome these limitations, Mahan *et al.*[6] developed, in 1993, a novel genetic system, which they termed '*in vivo* expression technology' (IVET). What made this strategy unique was their use of simple laboratory techniques to randomly choose genes (or only their promoters) from virulent pathogens, induce their expression within the host (i.e. *in vivo*), and then compare this with their expression *in vitro*. This then allowed them to select for those genes which were transcribed better or exclusively so only in the presence of the host environment. This, in fact, was the first attempt to study the induction of pathogenesis-related genes in the actual host environment conducive to disease development.

The first step in the technique consisted of the selection of a bacterial cell with a mutation in a biosynthetic pathway gene that greatly attenuated growth *in vivo*. For this purpose, Mahan *et al.* used a purine auxotroph (*purA*) of *Salmonella typhimurium*. Growth of such a strain *in vivo* was possible only if a wild-type *purA* gene could be introduced into it. A plasmid (pIVET1) was constructed by inserting a fusion of the wild-type *purA* gene with a marker gene (in this case, *lacZY*) into a broad host-range suicide vector (pGP704). The chromosomal DNA of *S. typhimurium* was fragmented by a specific restriction enzyme and cloned into the plasmid pIVET, 5′ to the *purA-lacZY* operon, thus giving rise to a transcriptional fusion where a *S. typhimurium* promoter would drive the transcription of a wild-type *purA-lacZY* fusion. A pool of pIVET1 containing these transcriptional fusion constructs was introduced into purine-deficient *S. typhimurium* cells.

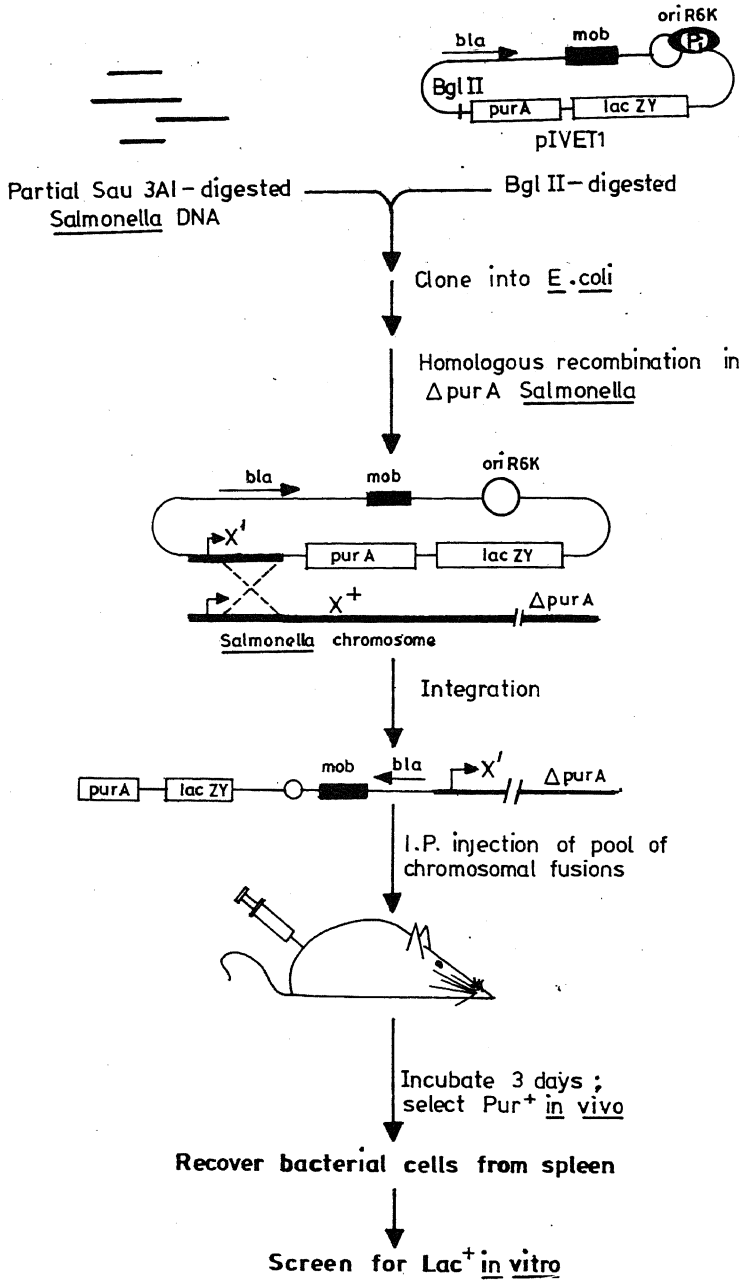


Figure 1. A schematic representation of the technique for the selection of *ivi* genes from *Salmonella*. Sau 3AI, Bgl II: restriction endonucleases (from Mahan *et al.*[5]).

Selection for ampicillin-resistant bacterial cells now enabled the identification of those in which the pIVET1 had integrated into the bacterial chromosome by homologous recombination with the cloned *S. typhimurium* DNA present on the plasmid (Figure 1). Thus, in these cells, the level of *purA* expression required to supplement the parental purine auxotrophy should logically correspond to the level of *lacZ* expression that would, in turn, give a Lac^+ phenotype. Only 33% of the cells grown *in vitro* showed Lac^+ and the corresponding $PurA^+$ phenotype. A pool of these recombinant bacteria were injected intraperitonially into a BALB/c mouse. The surviving bacterial cells were isolated from the spleen after three days and this procedure repeated once more. The cells that could overcome the purine deficiency due to the presence of the integrated *purA* gene were expected to survive and multiply in the mouse. The cells recovered from the spleen indeed showed an increase in their number (86%) compared to that in the initial inoculum. However, there were also a group of cells, though very small in number (5%), which showed a Lac^- phenotype in artificial medium. The question that naturally arises here is how could these cells manage to survive in spite of a very low probability of doing so. A reasonable explanation could be that the promoter that drove the *purA-lacZ* fusion gene in these cells could be turned on only *in vivo* and not under laboratory conditions. Therefore, the genes that were regulated by these promoters under natural conditions were aptly termed as '*in vivo*-induced' (*ivi*) genes. This phenomenon was confirmed by assaying β-galactosidase activity in the bacterial cells recovered from the spleen and comparing it with that of the same strain grown overnight in the laboratory, on enriched medium. In all the cases tested, the activity was actually 40 to 1000-fold greater in the cells grown *in vivo*.

The next step included the selection, cloning and sequencing of some of the *ivi* genes identified by the above technique. Sequence analysis could classify 15 of these genes into 5 different classes. Two of these turned out to be previously unknown genes with no homology to sequences in an established gene bank. These genes had probably not been identified earlier due to their inability to express *in vitro*. The other 4 genes included a *carAB* operon, a *pheST himA* operon and a *rfb* operon. The former consists of genes coding for two subunits of carbamoylphosphate synthase, an enzyme involved in arginine and pyrimidine biosynthesis. The low availability of pyrimidine in the animal tissue was obviously the force driving the induction of these genes. The IVET, in this case at least, helped in the identification of genes involved in intermediary metabolism and whose induction appeared to be necessary for the survival of the pathogen in the host tissues. The *pheST himA* operon is known to encode two subunits of phenylalanine tRNA synthetase and one subunit of the integration host factor (IHF). The latter is a DNA-binding protein required for essential functions like DNA replication, gene regulation and site-specific DNA recombination. The IHF has also

been implicated in regulating the expression of typel pilin (a protein conferring on *S. typhimurium* the ability to prevent clearance from animal tissues). A third fusion, the *rfb* operon, encodes approximately 20 genes involved in the synthesis of O-antigen (the outermost layer of lipopolysaccharide of *S. typhimurium*).

The ultimate goal of the experiments done by Mahan *et al.* was to identify the *ivi* genes involved in pathogenesis, which were otherwise not being picked up by the traditional techniques. Mutant strains, defective in *ivi* expression were then constructed to test their overall contribution to pathogenesis. These mutants, as expected, showed significantly less expression of the *lacZ* gene. This was also accompanied by a marked decrease in virulence as observed by a 2×10^4-fold increase in the lethal dose. Moreover, virulence was so diminished that mice could even be immunized with these mutant strains and rendered resistant to infection by *S. typhimurium*, administered orally.

Many other pathogenic bacteria are known to be attenuated by purine auxotrophy[7–11]. This technology can therefore be useful in studying a variety of host–pathogen systems. Various other biosynthetic genes can also be used for this selection scheme. One such example is a *thyA-lacZY* fusion construct which could be used in bacterial strains with thymine auxotrophy resulting in their attenuation. The advantage of this system is the easy selectability of thyA$^-$ mutants with trimethoprim.

A drawback of the IVET was its applicability only in host–pathogen systems where attenuating auxotrophies prevailed. However, earlier this year, Mahan *et al.* expanded the scope of this technology by developing a transcriptional fusion vector using antibiotic resistance as the basis for selection in host tissues[12]. This new vector contains a fusion of promoterless chloramphenicol acetyltransferase (*cat*) and *lacZY* genes. BALB/c mice were infected with a pool of *S. typhimurium* clones carrying random cat–lac transcriptional fusions and were subsequently given intraperitoneal injections of chloramphenicol. The strains which survived the passage through the animals expressed the *cat* gene and were then screened for those that had low-level *lacZY* expression on laboratory medium. These strains would carry fusion constructs containing specific *ivi* genes. A major advantage of the antibiotic-based IVET selection is that it does not require the isolation and characterization of an attenuated mutant strain with a genetically defined nutritionally-deficient species as in the original IVET protocol[6]. Thus, the antibiotic-based IVET selection should also be applicable to fastidious organisms where genetic markers, suitable for nutritional IVET schemes, are not readily available. Subtractable hybridization could be one way of identifying or estimating the expression of *ivi* genes of these microorganisms[13]. But the extremely low quantity of bacterial RNA produced in the host tissues could pose a limitation in achieving this.

One of the most challenging environments that invading bacteria encounter within the host is the host macrophages which eliminate the pathogens in the majority of cases. *S. typhimurium*, however, is an exception in that it has evolved mechanisms to survive within the inhospitable environment of the phagolysosomes. Genes of this organism, which are specifically induced in cultured macrophages, could also be isolated by the IVET. Characterization of the *ivi* genes of *S. typhimurium* in macrophages would help in tracing the unique mechanism by which this pathogenic species survives the unfavourable host environment.

With the introduction of antibiotic-based selection of *ivi* genes, the IVET can be extended to a broader host/host–pathogen range, thus opening up the possibility of studying a larger number of genes involved in various metabolic pathways in greater detail. Such studies may give us an insight into the underlying mechanisms regulating the development of different diseases and thus provide clues for the design of specific drugs against them.

1. Fields, P. I., Swanson, R. W., Hairdaris, C. G. and Heffron, F., *Proc. Natl. Acad. Sci. USA*, 1986, **83**, 5189–5193.
2. Fields, P. I., Groisman, E. A. and Heffron, F., *Science*, 1989, **243**, 1059–1062.
3. Isberg, R. R. and Falkow, S., *Nature*, 1985, **317**, 262–264.
4. Taylor, R. K., Miller, V. L., Furlong, D. D. and Mekalanos, J. J., *Proc. Natl. Acad. Sci. USA*, 1987, **84**, 2833–2837.
5. Miller, S. I., Kukral, A. M. and Mekalanos, J. J., *Proc. Natl. Acad. Sci. USA*, 1989, **86**, 5054–5058.
6. Mahan, M. J., Slauch, J. M. and Mekalanos, J. J., *Science*, 1993, **259**, 686–688.
7. McFarland, W. C. and Stocker, B. A., *J. Microbiol. Pathol.*, 1987, **3**, 129.
8. Baselski, V. S., Upchurch, S. and Parker, C. D., *Infect. Immun.*, 1978, **22**, 181.
9. Ivanovics, G., Marjai, E. and Dobozy, A., *J. Bacteriol.*, 1968, **85**, 147.
10. Levine, H. B. and Maurer, R. L., *J. Immunol.*, 1958, **81**, 433.
11. Straley, S. C. and Harmon, P. A., *Infect. Immun.*, 1984, **45**, 649.
12. Mahan, M. J., Tobias, J. W., Slauch, J. M., Hanna, P. C., Collier, R. J. and Mekalanos, J. J., *Proc. Natl. Acad. Sci. USA*, 1995, **92**, 669–673.
13. Plum, G. and Clark-Curtiss, J. E., *Infect. Immun.*, 1994, **62**, 476–483.

ACKNOWLEDGEMENT. I thank N. V. Joshi for attracting my attention to the recent developments in this area of research.

Kakoli Mukhopadhyay is in the Department of Crop Physiology, University of Agricultural Sciences, GKVK, Bangalore 560 065, India.

COMMENTARY

Are herbaria redundant?

R. R. Rao

In the current era of biotechnology and molecular biology the classical subjects like the taxonomy and herbarium have witnessed a great debacle. Unless immediate measures are taken to restore the herbaria and associated taxonomic activity, the entire biological community would soon end up in a state of confusion with regard to their own taxa of investigation and may have to knock the doors of scientists outside the country for seeking basic information on identity, nomenclature and even distribution of those taxa which occur at their door steps. While cautioning about such an ensuing serious problem I intend to outline here the paramount role of herbaria in allround development of botanical sciences including modern biology. This should also call for a dialogue between the classical biologists and experimentalists as to whether herbaria are redundant or they need to be maintained with added significance.

What is a herbarium?

Herbarium is a place where plants collected from far and wide are preserved in pressed and dried condition and are kept in pigeon holes of almirahs according to some accepted system. It is a great filing system for information about plants primarily in the form of actual specimens and secondarily in the form of recorded notes on labels attached on the sheets. In other words, herbarium is a vast reservoir of facts about plants.

Functions or utility

In order to wash out the wrong notion among a section of modern biologists it is necessary to highlight the diverse functions of a modern herbarium and prove how it acts as a connecting link between various disciplines. The following are a few important functions of a herbarium.

* Conservatory of material and data
* As a storehouse of collections including the valuable type specimens the herbaria greatly aid in all kinds of taxonomic researches
* Aids in teaching and in all biological researches
* Serves as a fundamental resource for identification of all plants of the globe
* Serves as source for collection of biodiversity information (to capture biodiversity information from herbaria is at least 10 times more cost efficient than to collect new data from field). Most estimates on global biodiversity today is based on herbarium collections only
* Aids in biodiversity monitoring; carrying out scrutiny of herbarium collections to obtain quantitative baseline data on the distribution and abundance of keystone species is essential for all monitoring programmes
* Aids in assessment of conservation status of a taxon (helps in cataloguing of rare and endangered species for conservation programme)
* Serves as a repository of historic collections and at times acts as the only record of past vegetation
* Serves as a source for search of new genetic material for improvement of our cultivated stock
* Aids in assessment and cataloguing of all species of economic potential; these may include commercial species, medicinal herbs, food plants and so on
* Serves as a source of material for investigations on anatomy, morphology, ecology, chemistry, molecular biology, pharmacognosy and for environmental impact assessments
* Seeds of herbarium specimens can be used to resurrect species extinct in the wild using modern technologies
* Herbarium, like a library, is a service institute; solves all queries about plants from scientists, administrators, lawyers, corporate personnel, industrialists, environmentalists and foresters
* Helps in development of computer databases on plants and maintains active links to international network of systematic resources and electronic knowledge base.

Thus, it is clear that herbaria contribute to the development of all biological disciplines. But today herbaria are completely ignored by the so-called modern biologists who have least knowledge of the significance of a herbarium. Although herbarium is a facility of a data bank on plants to be used by all biologists only a handful of surviving taxonomists are putting their efforts to nourish and maintain the fragile heritage of herbaria in the country. This has resulted in erroneous conclusion that herbaria are the establishments solely of taxonomists. Rather, herbaria are national facilities and their maintenance should be a national responsibility.

Some herbaria developed over several decades of efforts of taxonomists are today at the verge of collapse due to wrong impression among the ruling biologists that herbaria are merely a storehouse of collections of dead plants, which cannot contribute to the national development nor can generate funds for research forgetting that herbaria are simply a facility of a databank on plants from which all biologists draw their basic informations directly or indirectly about the plant species on which they carry out all advanced researches.

The significance of voucher specimens, on which varieties of researches are carried out, need not be elaborated here. It is only the herbaria which maintain such voucher specimens for posterity. Time and again the necessity of depositing the voucher specimens in some recognized herbaria has been stressed. But today this aspect is so much ignored that many modern biologists carry out outstanding researches on taxa but without maintaining voucher specimens nor do they feel the necessity. In the absence of voucher specimens vast amount of new data/new discoveries on a particular plant species quite likely may not pertain to the plant species studied. Even if the species is well known it may contain a number of ecotypes, cytotypes, varieties or strains and hence it is absolutely necessary to maintain the voucher specimens in any

herbarium to allow subsequent researchers to evaluate statements made in the literature about the characteristics of a given taxon. The editors of various scientific journals should also ensure that all authors cite the accession number/numbers of their voucher specimens and the herbaria where deposited. While this enhances the value of scientific findings, at the same time it strengthens the justification of the importance of the herbaria.

Thus, herbaria are helpful to both classical and experimental biologists not only for obtaining the basic data on plants but in also serving as a caretaker of the voucher specimens on which tremendous amount of literature accumulates. In view of this, are we unjustified in demanding for allocation of some percentage of funds from all research projects as service charge for development and maintenance of herbaria? It is also certainly unwise to score out the herbaria and taxonomic researches as establishments of no economic gains as the present trend goes.

National herbaria like the Central National Herbarium, Herbarium of the Forest Research Institute, Dehra Dun, and the Herbarium of the National Botanical Research Institute, Lucknow are critically endangered due to lack of sufficient trained manpower, facility and even due recognition by the so-called experimentalists. With the overgrowth of several modern disciplines the importance of herbaria has faded, resulting in such a damage that would be felt in the immediate coming years. Khoshoo[1] has rightly pointed out that taxonomists are a 'vanishing tribe' among the biologists and are greatly overshadowed by the so-called and more often second-rate biotechnologists and environmentalists.

Herbaria require large buildings and staff for curation of vast collections, laboratories for associated researches and funds for continuous explorations for enriching the herbaria, particularly in a developing country like India where the coverage of holdings is incomplete. The knowledge of variation within species is another aspect which is very limited. But the increasing financial squeeze and thoughtless prioritization of research programmes have greatly affected the overall health of the herbaria in the country. Often comments from scientists state that the massive collections in the herbaria be dispensed with and only few sample collections and type specimens be maintained[2]. There is therefore, an urgent need to educate the policy makers, the experimental biologists, and other key persons who matter much for development of herbaria in the country about the 'essentiality' of a herbarium.

A rational economic basis for maintenance and furtherance of herbarium research must consider the fact that botanists with all their concerted efforts have known only 1/10th of what exists in our tropical forests and still less is known of the economic utility of those species which are recorded. Infraspecific biodiversity and population variation of the rich tropical flora are little understood. But certainly maintaining and enriching a herbarium is expensive. A strict monitoring of the quality of the incoming collections and zones from where they come is essential in order to maintain the quality and not the size of the herbarium. More than one herbaria if any, within a city or town (e.g. Dehra Dun, Lucknow) can be considered for merging in view of the increasing cost of maintenance, manpower, space on buildings, etc., rather than 'killing' a herbarium due to wrong policies and apathy towards such classical subjects upon which many future solutions depend.

1. Khoshoo, T. N., *Curr. Sci.*, 1995, **69**, 19–17.
2. Clifford, H. T., Rogers, R. W. and Dettmann, M. E., *Nature*, 1990, **346**, 602.

R. R. Rao is in the National Botanical Research Institute, Lucknow 226 001, India.

SCIENTIFIC CORRESPONDENCE

The academy fellowship problem

The problem

A science academy elects new fellows every year. In order to improve the overall quality of its fellowship, it decides to impose a new criterion that every new fellow elected shall be better than the median level of the existing population of fellows*. How fast will the quality of the fellowship improve?

*This criterion was proposed for the Indian Academy of Sciences by Professor V. Radhakrishnan in the mid-seventies.

A toy model

The above problem needs to be recast in a quantitative form. It is difficult to quantify 'quality' of a scientist but suppose in a toy model we use a variable x to measure it on a scale ranging from 0 to 1, with 1 being the mark of perfection. Let us measure the time in years with $t = 0$ denoting a starting year when this criterion is announced to be implemented in all subsequent years. Let $f_n(x)$ denote the cumulative quality distribution of fellowship in the year $t = n$, $n = 0, 1, 2, \ldots$ In the starting year the total number of fellows was N_0, say and suppose that by statutes, every year N fellows must be elected.

Denote by M_n the median of the distribution $f_n(x)$. Thus

$$f_n(M_n) = \frac{1}{2} f_n(1) = \frac{1}{2}(N_0 + nN), \qquad (1)$$

and while electing in year n, care is taken that all new additions shall have the quality parameter $x > M_{n-1}$. Improvement in quality will then be indicated by how the median M_n steadily increases in value year by year.

We still need to specify the quality distribution of the pool of scientists

from which the new fellows are chosen. We denote by $g(x)$ the cumulative probability that a scientist selected at random from this pool will not have the quality parameter exceeding x. Thus

$$g(0) = 0, \qquad g(1) = 1. \tag{2}$$

In this toy model we have ignored the reduction in the total population of fellows due to death or resignation (or even eviction).

With these specifications we can now pose the question more precisely: Given the functions $f_0(x)$ and $g(x)$, how does M_n increase with n?

Solution

First note that for $t = n$, the median criterion implies

$$f_n(x) = f_{n-1}(x) \text{ for } x < M_{n-1}. \tag{3}$$

For $x \geq M_{n-1}$, there will be addition of N fellows whose distribution will be assumed to be proportional to the function $g(x)$ over $M_{n-1} \leq x \leq 1$. Therefore, we have for $x \geq M_{n-1}$

$$f_n(x) = f_{n-1}(x) + N\,\frac{g(x) - g(M_{n-1})}{1 - g(M_{n-1})}. \tag{4}$$

Since for the distribution $f_n(x)$ there are $\frac{1}{2}(N_0 + nN)$ members with $x < M_n$, these comprise of the $\frac{1}{2}\{N_0 + (n-1)N\}$ members with $x < M_{n-1}$ together with the number $f_{n-1}(M_n) - f_{n-1}(M_{n-1})$ from the earlier distribution and the new addition as per (4) above. Carrying out this book-keeping we have

$$f_{n-1}(M_n) - f_{n-1}(M_{n-1}) + N\,\frac{g(M_n) - g(M_{n-1})}{1 - g(M_{n-1})} = \frac{1}{2}N. \tag{5}$$

Next we consider the number $f_{n-1}(M_n) - f_{n-1}(M_{n-1})$ which has arisen from steady addition to $f_0(M_n) - f_0(M_{n-1})$ over all the preceding years following the rule (4). Therefore,

$$f_{n-1}(M_n) - f_{n-1}(M_{n-1}) = N \sum_{r=1}^{n-1} \frac{g(M_n) - g(M_{n-1})}{1 - g(M_{r-1})} + f_0(M_n) - f_0(M_{n-1}). \tag{6}$$

From (5) and (6) we have the final relation

$$f_0(M_n) - f_0(M_{n-1}) + N\{g(M_n) - g(M_{n-1})\} \times \sum_{r=1}^{n} \frac{1}{1 - g(M_{r-1})} = \frac{1}{2}N. \tag{7}$$

This iterative relation determines the sequence $\{M_n\}$ in a step by step fashion starting with M_0, since for determining M_n we have all other quantities known from the preceding application of this relation.

A simple example

We will illustrate the above solution with a simple example in which both the initial distribution and the general population distribution are uniform with respect to the attribute x. In this case

$$f_0(x) = N_0 x, \qquad g(x) = x. \tag{8}$$

and the relation (7) gives

$$M_n = M_{n-1} + \frac{1}{2\sum_{r=1}^{n} \frac{1}{1 - M_{r-1}} + \frac{2N_0}{N}}. \tag{9}$$

Even with this simple example we cannot find a solution in closed form, but Figure 1 illustrates the march of M_n for $0 \leq n \leq 25$, $N_0 = 500$ and $N = 20$ obtained by numerical methods. The median value rises from 0.5 to 0.75 over twenty-five years.

The continuum version

The solution, however, becomes tractable even for the general case if we modify the problem by having a continuous input of fellows at all times instead of the discrete annual input. A simple analysis shows that the relation (7) is changed to an integro-differential equation

$$\left(\frac{dM}{dt}\right)^{-1} = \frac{2}{N}\left.\frac{\partial f(x,t)}{\partial x}\right|_{x=M(t),\,t_0} + 2\int_0^t \frac{g'(M(t_1))\,dt_1}{\{1 - g(M(t_1))\}}, \tag{10}$$

where the prime denotes derivative with respect to the argument. Here the rate of

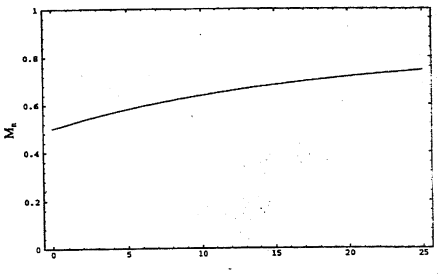


Figure 1. The figure shows how the median value of the fellowship quality improves with successive elections over 25 years in the simple toy model.

fellowship input is N, that is $N\delta t$ to be added in a time interval δt. Also, the fellowship distribution at time t is now denoted by $f(x, t)$ and its median by $M(t)$.

The equation (10) is solvable in closed form and in the general case we have

$$\frac{dt}{dM} = \frac{2}{\{1 - g(M)\}^2} \times \left[\int_{M(0)}^{M(t)} \alpha'(M)\{1 - g(M)\}^2\,dM + A\right], \tag{11}$$

$$\alpha(M) = \frac{1}{N}\left.\frac{\partial f(x,t)}{\partial t}\right|_{x=M,\,t=0}, \tag{12}$$

where A is a constant of integration. For the simple case of uniform distributions discussed earlier the solution is very simple:

$$M(t) = \frac{1}{2}\cdot\frac{1 + (2tN/N_0)}{1 + (tN/N_0)}, \tag{13}$$

and it fits the curve of Figure 1 very well. In fact an approximate discrete solution based on equation (13), viz.

$$M_n = 1 - \frac{1}{2}(1 + nN/N_0)^{-1} \tag{14}$$

satisfies the iterative relation (7) with a good approximation.

JAYANT V. NARLIKAR

Inter-University Centre for Astronomy and Astrophysics,
Post Bag 4, Ganeshkhind,
Pune 411 007, India

Supplementary comments on 'endemic, rare and threatened flowering plants of South India'

This note is to supplement Meher Homji's comments (*Curr. Sci.*, 1995, **68**, 1083–1084), on the article by Ranjit Daniels *et al.* (*Curr. Sci.*, 1995, **68**, 493–495).

Coining a term 'pseudoendemic' for a plant species like *Piper barberi* occurring in the southernmost parts of the W. Ghats in Kanyakumari District and also in a very closely adjacent and contiguous area in Kulathupuzha is rather misleading. Plants can be either 'endemic' or wide species. There is no 'pseudo' category in the plant or animal world. There are well-defined categories of endemics such as palaeoendemic, neoendemic, patroendemic, schizoendemic and apoendemic. It is better not to add new phraseology like 'pseudoendemic'.

As mentioned by Meher Homji, conjectural propositions on extinction phenomena based on insufficient data, that too from a small politico-geographical area, (Tamil Nadu!) are not appropriate. Biological extinction and man-made extinction must be conceived in their true perspectives. Biological extinctions involve the processes of 'way out' of species through millions of years of speciation and extinction in time space continuum. Unfortunately the current crisis of extinctions of biota is steered by anthropogenic intervention and habitat loss.

M. P. NAYAR

Tropical Botanic Garden and Research Institute,
Pacha-Palode P.O.,
Thiruvananthapuram 695 562, India

R. J. Ranjit Daniels et al.'s reply:

We thank both V. M. Meher-Homji and M. P. Nayar for their interesting comments on our paper (*Curr. Sci.*, 1995, **68**, 493–495). In fact, the primary purpose of the paper was to invoke some discussion.

Regarding the issues raised by Meher-Homji (*Curr. Sci.*, 1995, **68**, 1083–1084) and that subsequently quoted by Nayar, we admit that Tamil Nadu is not 'south India'. However, we have just chosen to treat the plants of Tamil Nadu listed in the *Red Data Books* as a subset of those reported from the 'Madras Presidency' by Gamble. Madras Presidency, which included, at least in part, the four south Indian states, viz. Kerala, Tamil Nadu, Karnataka and Andhra Pradesh is in our opinion and geographically a fairly good representation of south India. We therefore feel that it is statistically appropriate to work with the available data on small subsets for drawing more general inferences.

On the issue of 'biological extinction and human-induced extinction', it is not clear to us how the distinction can be easily made. We are only highlighting the fact that certain species are 'more extinction-prone' than the others. We have also illustrated some of the possible reasons that make such species extinction-prone. Species on their natural 'way out' are as vulnerable to extinction as those which are being 'led' by humans. Whether we pay more importance to one at the expense of the other becomes more of a philosophical issue.

Finally, regarding the 'pseudoendemics' and the contention that there is nothing 'pseudo' in the natural world, we have an explanation as well. One has to admit the fact that all the status that we humans attribute to living organisms are 'relative' and at any stage be claimed as false. While it cannot be rationally argued as to why we should not use such terms as pseudo, it is easier to bring to the attention of the contender the fact that this term was not 'coined' by us. Although while writing this paper we thought we were introducing a new term to describe 'endemic' species, we discovered later that this term has been in use since 1990 after its first introduction by Nelson *et al.* (*Nature*, **345**, 714–716). The term was, as much as we have done, used to recognize 'endemic' species whose full range has not been sampled due to inadequate or uneven surveys.

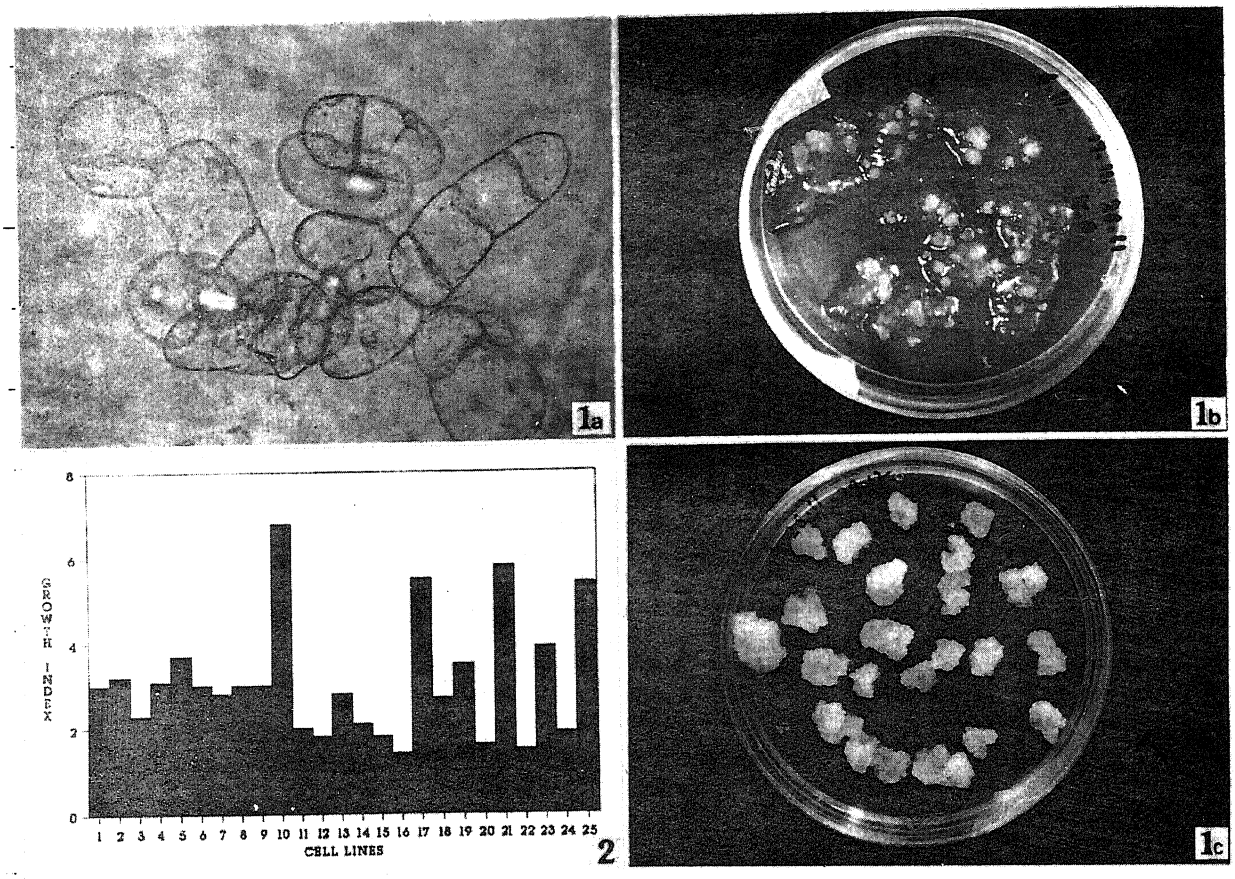
R. J. RANJIT DANIELS
N. ANIL KUMAR
M. JAYANTHI

M. S. Swaminathan Research Foundation,
3rd Cross Street,
Taramani Institutional Area,
Madras 600 013, India

A fast-growing cell line of *Taxus baccata* L. (Himalayan yew) as a potential source of taxol precursor

Taxol, a diterpene with exceptional anticancerous activity, is presently isolated from bark of Pacific yew, *Taxus brevifolia*[1]. It occurs as a very minor component in several species of *Taxus*[2]. Continued bark harvesting for taxol does not appear to be an adequate source to meet its demand[3]. There is a growing interest in alternative methods for taxol production such as from tissue culture of *Taxus* sp.[4–6]. The Himalayan yew, *T. baccata*, has been reported[7] to contain taxol and 10-deacetyl baccatin III, a precursor for synthesis of taxol. Here we report selection and cloning of fast-growing cell lines of Himalayan *T. baccata* for the production of taxanes.

Callus cultures were initiated from needle leaf of mature trees of *T. baccata* (Linn.) collected from Darjeeling area during July–October 1993. Callus was induced on explants on B_5 medium[8] supplemented with 2 mg/l 2,4-D, 0.1 mg/l kinetin and 0.1% casein hydrolysate, as reported for other species of *Taxus*[4,6]. Culture growth was main-



Figures 1 and 2. 1*a*, Early exponential phase cells of *T. baccata* suspension, *b*, visible colonies following plating after 21 days and *c*, variability in growth rates of cell colonies after 8 weeks of plating. **2.** Comparison of growth index of 25 established cell lines of *T. baccata*.

Table 1. Taxol and 10-deacetyl baccatin III content in four fast growing cell lines of Himalayan *T. baccata* after 6 months of culture

		Taxol		10-deacetyl baccatin III	
Cell line	GI	*Cell suspension (mg/100 g)	Media (µg/l)	*Cell suspension (mg/100 g)	Media (µg/l)
Kp 510	6.8	3 ± 0.5	10	242 ± 15	620
Kp 517	5.5	10 ± 1.5	33	177 ± 21	370
Kp 535	5.8	6 ± 0.5	15	160 ± 11	310
Kp 551	5.4	7 ± 1.2	Not detectable	168 ± 10	112

*Values represent mean ± standard error expressed on a dry matter basis.

tained on the same medium and was subcultured at 4-week intervals for six months. Suspension cultures were initiated by transferring callus tissues (2 g fresh wt) into 50 ml B_5 medium containing 2,4-D (0.2 mg/l) and kinetin (0.1 mg/l) in 250 ml Erlenmeyer flasks. The cultures were incubated on a rotary shaker (125 rpm) at 25 ± 2°C under continuous dark and were transferred every 15 days. Cell suspension cultures were maintained by inoculating aliquotes of filtered cultures (nylon mesh 0.5 mm pore size) equivalent to 10 mg dry wt (100°C/5 h) per ml of fresh medium containing 1 mg/l 2,4-D and 0.05 mg/l kinetin. Cultures were regularly passed through nylon sieve (0.5 mm) up to 6 months to obtain a fine cell suspension culture.

For selection of fast-growing cell lines, 3-week-old cell suspension cultures (GI 1.8 to 2.5) were centrifuged at 2000 rpm for 5 min, the liquid medium discarded and fresh medium added. Appropriate aliquotes were mixed in petri dishes with MS agar medium, held liquid at 40°C to give a final cell density of 5×10^4 cells/ml. The final agar concentration was 0.6%. The petri plates were incubated under continuous dark at 25 ± 2°C for 45 days. Cell division started within 7 days (Figure 1 *a*, *b*) and visible colonies (Figure 1 *c*) appeared on agar plates within 21 days of culture. Variability in cell growth rate was observed in different cell lines (Figure 1 *d*) initiated from the same explant. Growth index was calculated as (final wt – initial wt)/initial wt. Using this procedure, it was possible to select several cell lines showing faster growth. One major problem associated with *Taxus* cell suspension culture is the production of red coloured exudates (presumably phenolics) which effected cell growth. By following selection procedure as

described above, creamish-white cell suspension cultures could be established with minimum exudation of phenolics. These cell lines have been maintained for over 12 months in culture.

Taxanes in cell suspension cultures were measured using previously described methods[6]. For analysis of taxanes in cell suspensions and cultured media, cultures were harvested after 3 weeks of each passage, fresh weight of cell cultures recorded and the material lyophilized and extracted with chloroform (20 ml/g cells). The liquid medium was shaken with chloroform (100 ml/l). The chloroform extracts were evaporated to dryness, suspended in 5 ml methanol (HPLC grade) and filtered (0.2 μm) prior to HPLC analysis. HPLC analysis of taxol and 10-deacetyl baccatin III was performed in a Shimadzu liquid chromatograph (LC 10 AD) employing a reverse phase ODS–Supelcosil C-18 column (25 cm × 4.6 mm id), a Supelco guard column (Pelliguard™ LC-18 kit with 2 cm × 4.6 mm cartridge) and isocratic elution with acetonitrile:water:methanol (30:30:40). The flow rate was 1 ml/min and all chromatographs were plotted at 227 nm, using a UV detector (SPD 10A). A 0.25 mg/ml solution of standard samples of taxanes in MeOH was prepared and filtered through 0.2 μm filter prior to HPLC analysis. Identity of taxol and 10-deacetyl baccatin III was confirmed by retention time, spiking with standard and on the basis of spectral properties[9,10].

Extreme variability in cell growth rate was observed (Figure 2) in cell lines established by plating. Selection and recloning of faster-growing colonies led to establishment of four cell lines with rapid doubling time. Such variability in growth rates of cell lines has been reported in cell cultures of *T. brevifolia*[7,11].

Suspension cultures were established from the four cell lines (GI 5.5–6.8) and taxol and 10-deacetyl baccatin III were detected in cell suspension and culture media. However, taxol content was low and ranged from 0.003% to 0.007% in the four cell lines. The content of 10-deacetyl baccatin III varied from 0.162% to 0.212% (Table 1). Low levels of taxanes were also exuded in the liquid culture medium. The results are particularly significant since total chemical synthesis of taxol has not yet been achieved and an efficient semi-synthetic method starting from 10-deacetyl baccatin III has been reported[12]. Thus cell suspension culture of Himalayan *T. baccata* may be used as a source of precursor for synthesis of taxol.

1. Wani, M. C., Taylor, H. L., Wall, M. E., Coggon, P. and McPhail, A. T., *J. Am. Chem. Soc.*, 1971, **93**, 2325.
2. Wiltherup, K. M., Look, S. A., Stasko, M. W., Ghizori, T. J. and Muschik, G. M., *J. Nat. Prodts.*, 1990, **53**, 1249–1255.
3. Heinstein, P. F. and Chang, C. J., *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, 1994, **82**, 1247–1259.
4. Fett Neto, A. G., DiCosmo, F., Reynolds, W. F. and Sakata, K. O., *Bio/Technology*, 1992, **10**, 1572–1575.
5. Fetto Neto, A. G., Melanson, S. J., Sakata, K. O. and DiCosmo, F., *Bio/Technology*, 1993, **14**, 731–734.
6. Gibson, D. M., Ketchum, R. E. B., Vance, N. C. and Christen, A. A., *Plant Cell. Rep.*, 1993, **12**, 479–482.
7. Chattopadhyay, S. K., Tripathi, V. K., Thakur, R. S., Sharma, R. P. and Jain, S. P., *Indian J. Chem.*, 1994, **B34**, 409–411.
8. Gamborg, O. L., Miller, R. A. and Ojima, K., *Expo. Cell Res.*, 1968, **50**, 151–158.
9. Falzone, C. A., Benesi, A. B. and Lecomte, J. T. J., *Tetrahedron Lett.*, 1992, **33**, 1169–1172.
10. Miller, R. W., Powell, R. G., Smith, C. R., Arnold, E. and Clardy, E., *J. Org. Chem.*, 1981, **46**, 1469.
11. Wickremesinhe, E. R. M. and Arteca, R. N., *J. Plant Physiol.*, 1994, **144**, 183–188.
12. Denis, J. N., Greene, D., Guenard, F., Gueritte-Voegetem, F., Mangalal, L. and Potier, P., *J. Am. Chem. Soc.*, 1988, **1110**, 5917.

ACKNOWLEDGEMENTS. We thank the Drug Synthesis and Chemistry Branch, developmental Therapeutics Program Division of Cancer Treatment, National Cancer Institute, Maryland, USA for standard sample of taxol and Dr P. Potier, France for the sample of 10-deacetyl baccatin III.

SUMITA JHA
TIMIR B. JHA

Department of Botany,
Calcutta University,
35, B.C. Road,
Calcutta 700 019, India

Solar photovoltaic energy products: from semiconductor physics to advanced energy technology

Ashok Parthasarathi, D. Madhavan and V. K. Kaul

Physics to industry

There is a long road to be traversed between the original enunciation of a fundamental physical principle and the final realization of an application of that physical principle in the form of an industrial product useful to society. The bridging of this gap between physics and industry consists of a series of innovative steps ranging from perfecting laboratory scale processes to converting those processes and associated techniques through up-scaled, standardized production methods in a pilot plant, to large volume, cost-effective production on a commercially viable basis. This also includes the development of appropriate production equipment as needed. It is the collection of such knowledge, skills and equipment which constitutes the development of a complete technology package by which the physical principle is converted into an industrial technology to make a commercial product.

This paper presents such a case of 'Bridging the Gap' between semiconductor physics and the solar photovoltaic (SPV) industry. The case takes the form of the actual history of development and commercialization of SPV technology by our public sector company, Central Electronics Limited (CEL). CEL is today not only the nation's pioneer and largest manufacturer by far in the developing countries but the 5th largest manufacturer of single crystal silicon solar cells in the world.

Historical evolution of photovoltaics (PV)

The Photovoltaic (PV) Effect is the appearance of a voltage across a p–n junction or a rectifying contact as a result of its exposure to light. The light produces electron hole pairs. These pairs are separated by the inherent electric field across such junctions which causes a current to flow in the external circuit through the leads attached to the two sides of the junction.

Paper presented earlier in a more detailed form at the International Conference on Physics and Industrial Development: Bridging the Gap January 17–19, 1994.

Ashok Parthasarathi is in the Department of Scientific and Industrial Research, Ministry of Science & Technology, New Delhi, India. D. Madhavan and V. K. Kaul are at the Central Electronics Limited Sahibabad 201 010, India.

The Photovoltaic Effect was first discovered in an electrochemical cell by Becquerel[1] in 1839. The first solid state photovoltaic cell of selenium was made in 1877 by Adams and Day[2]. Although silicon solar cells based on natural (non-doped) junctions were made in 1941 (refs. 3, 4), it was only in 1954, that the first diffused p–n junction silicon solar cell was made by Pearson[5] and his group at Bell Telephone Laboratories, USA. It is that Pearson Cell which forms the basic building block of present-day solar photovoltaic energy generating systems.

SPV systems were first used commercially in 1957 as 'solar batteries' to power man-made earth satellites and this was the main application of solar cells till the mid–70's. However, it was the so-called 'oil crisis' of 1973 which spurred the development of technology for making solar cells, modules and complete SPV systems for *terrestrial applications*.

Solar PV systems are widely used to power various systems – both rural and industrial. Solar photovoltaic energy sources produce DC electricity directly from solar energy; they are modular in nature and their maintenance requirements are minimal; they are also absolutely non-polluting. Unlike other sources of electricity, the cost per unit of electrical energy generated does not depend drastically on the capacity of the photovoltaic energy generating unit installed. Therefore, small stand-alone SPV sources are just as viable economically as large and centralized ones, even at the present level of technology. All these advantages make solar photovoltaic energy sources eminently suitable for use in places where electricity is not available from conventional sources such as hydro, thermal or nuclear power generators and where diesel generators or electrochemical cells are not economical or feasible to use.

Single crystalline silicon technology is well understood and is mature due to its preponderant use in microelectronics. Hence, it is the principal raw material for making terrestrial solar cells as compared to other photoconductive semiconducting materials such as gallium arsenide.

The silicon PV chain

The silicon PV chain starting from the initial raw material polysilicon to the final SPV system is depicted in

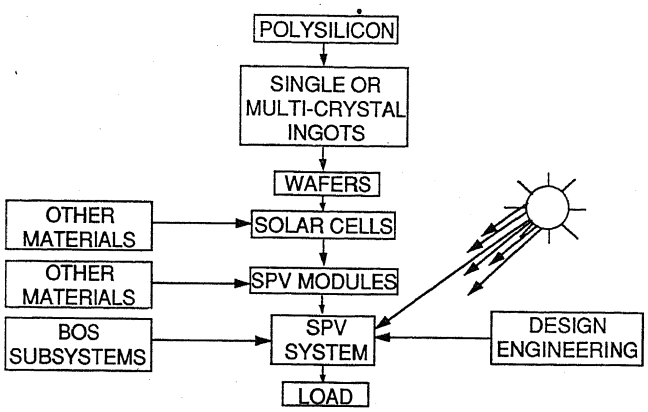


Figure 1. Schematic of silicon PV chain.

Figure 1. Polycrystalline silicon nuggets are the starting point in the PV chain. These are melted in a furnace and grown into single crystalline ingots. Such ingots are commercially grown by the standard Czochralski method upto a diameter of 100–150 mm. Then silicon wafers (thickness 200 to 400 microns) are sliced out of the ingots either by diamond impregnated internal diameter (ID) cutters or through multiple wire saws. The wafers are converted into solar cells through appropriate process technology using materials such as special chemicals, dopant elements e.g. phosphorus and boron and special conducting pastes of silver and aluminium compositions. These processes are drawn from the now well-known semiconductor device fabrication process, since the silicon solar cell is nothing but a large area semiconductor diode. A number of these cells are then interconnected and made into modules by means of basically a lamination process. This process involves the use (apart from the solar cells themselves) of other specialized materials such as tough tempered glass, soft polyester materials, metal frames and a junction box for extracting the electrical power. These SPV modules are the starting point for configuring a complete SPV system appropriate to a specific application. The final engineering of an SPV energy system is based on site and load-specific design including use of special computer software developed for that purpose. All the sub-systems other than the modules and the load (e.g. batteries, charge controllers/inverters) are termed the '*Balance of* the total Solar Photovoltaic *System*' (BOS).

The techno-economic performance of a solar cell or module is measured by its conversion efficiency (η). This is the percentage of the incident solar light energy which is converted by the cell/module into electrical energy. Today, n^+–p junction circular silicon solar cells of 100 mm diameter or square cross-section cells of 100 mm side have efficiencies in the range of 11%–14% in commercial volume production. The major factor presently preventing large-scale commercial exploitation of these power-generating devices is their high initial cost. Increase in the conversion efficiency has been recongnized as one of the key factors which would bring down the cost of photovoltaic power generation. To this end, researchers in laboratories and companies around the world are looking critically at the various loss mechanisms and attempting technological innovations to overcome the losses so as to approach the theoretical conversion efficiency limit of around 30% of n^+–p junction cells. In recent years, hand-made cells in the laboratory using hyperpure silicon wafers have shown efficiencies of 18%–19% for cells of 70 mm × 70 mm (45–50 cm^2)[6] and as high as 24% for small size cells (20 mm × 20 mm corresponding to 4 cm^2)[7]. The challenge is to achieve these higher efficiencies as *average* efficiencies of thousands of large area (100 mm × 100 mm) cells per day in commercial production in industrial plants and that too with commercial-grade CZ silicon wafers as the starting raw material.

SPV in India

Until 1975, activity in the field of SPV in India consisted only of R and D on solar cells in a few national laboratories. Development activity was directed only towards laboratory-scale solar cell fabrication, mainly oriented towards possible future space applications.

SPV in CEL

The activity at CEL in the area of solar photovoltaics extending over a decade and half, can be broadly broken up into the following macro-milestones:

1975–1980	:	R & D phase
1981–1985	:	Pilot plant operation
1986–1991	:	Industrial-scale technology proving and semi-commercial operations
1991 onwards	:	Commercial operation.

The development of solar cells for terrestrial applications was initiated at CEL following Government's decision, in 1975, to mount concerted efforts in this high-technology area. CEL has carried out extensive in-house R & D work spanning a decade for developing the complete technology for the manufacture of silicon solar cells and modules and designing, engineering and operating a pilot plant for production of such cells and modules based on the process technology and production engineering so developed. The activity also includes the development of a whole range of SPV systems and undertaking large-volume commercial production, supply, field installation and commissioning of such systems. Starting with processing of 38 mm diameter silicon wafers using vacuum metallization in 1978, CEL went through an evolu-

Table 1. Milestones of cell/module process technology development at CEL

Processing of 1" diameter solar cells	1977
First SPV module with aluminium substrate	1977
Liquid encapsulation of cells into modules	1977
Vacuum metallization of cells	1977
Processing of 2" diameter solar cells	1978
Processing of 3" diameter solar cells	1980
Introduction of expansion loop in cell interconnection	1981
Use of tempered glass in module instead of plate glass	1981
Introduction of glass substrate instead of aluminium for intrinsically safe modules	1982
Screen-printed metallization of cells instead of vacuum metallization	1983
Processing of 4" diameter solar cells	1984
First foldable module for expedition to Mount Everest	1984
Use of low-iron tempered glass for supertrate	1986
Tedlar–aluminium–tedlar substrate instead of aluminium	1986
Air-firing of cells after metallization	1988
Sirtl etch for edge isolation	1989
Processing of 4" pseudosquare solar cells	1989
Lead attachment to cells using automatic tabbing machine	1991
Automated chemical processing station	1993

tionary development process in terms of both different sizes of cells and the whole range of process technology for making them. The milestones of development of solar cells and modules showing the evolution of the process and production technology at CEL is given in Table 1.

Solar cell process technology

The solar cell production process flow chart is shown in Figure 2. The process flow chart corresponds to the Screen Printed Technology (SPT) of making silicon solar cells, which is the dominant technology currently in use the world over and also the technology developed and in commercial use at CEL. A perspective of CEL's modern solar cell process area is shown in Figure 3. The individual process steps in the SPT are as follows:

Surface preparation

The raw silicon wafers sometimes have oil or grease on their surface, due to adhesion of the cutting coolant used at the time the wafers are sliced out of the silicon ingots. Therefore, it is essential to prepare the silicon surface properly for subsequent polishing and texturization by cleaning its surface with suitable non-ionic detergents.

Chemical polishing and texturization

The raw silicon wafers also have saw marks which are likely to affect the performance of the solar cells made from such wafers. So, these saw marks need to be removed. This is done by alkaline polishing of the wafers. Further, in order to absorb more incident light, the polished wafers are subjected to texturization by preferential etching with low concentration alkaline solutions.

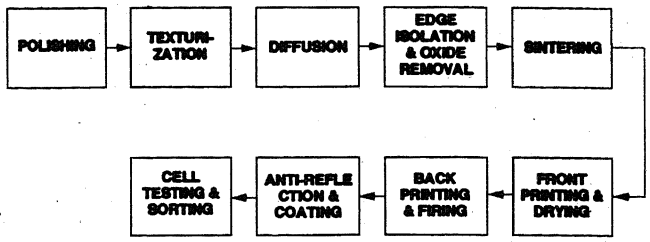


Figure 2. Solar cell process flow chart.

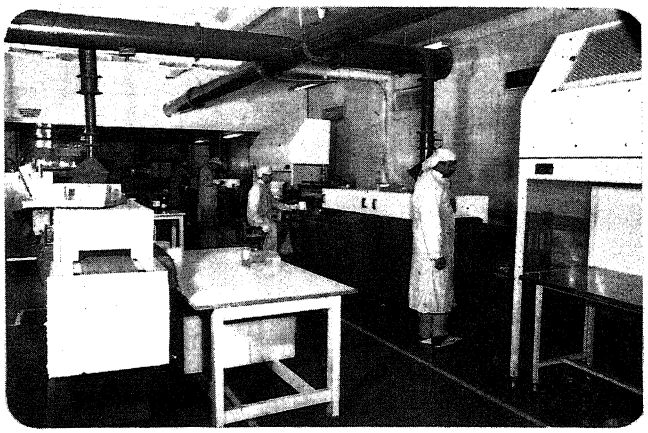


Figure 3. A perspective view of CEL's modern solar cell process area.



Figure 4. Indigenous automated chemical station.

All these chemical process steps are now being carried out using an indigenously designed and fabricated automated chemical station. The chemical station has been designed and fabricated in association with the Bhabha Atomic Research Centre (BARC), Bombay (see Figure 4).

Diffusion

This is the process step in which the n–p junction is formed on the silicon substrate. This process is carried

out using $POCl_3$ and N_2 and O_2 gases in a semi-automated diffusion furnace incorporating a microprocessor-based gas control and sequencing system.

Edge isolation

This process step is for the isolation of the junction at the periphery of the diffused silicon wafers. It is achieved either by chemical etching, plasma etching or laser scribing.

Back Surface Field

One of the limiting parameters of solar cell performance is the open circuit voltage which is limited due to the particular method of junction formation. This parameter (*i.e.* open circuit voltage) can be further enhanced by introducing another heavily doped p–type layer popularly known as back surface field. Screen-printed aluminium metallization is undertaken on the back of the wafer as the key step in this process. The introduction of the BSF enhances the overall efficiency of the solar cell.

Screen printing

For providing the contacts for carrier collection and interconnects both sides of the silicon, suitable conducting pastes are screen-printed and fired. This step is carried out by two semi-automatic integrated screen printers – one for screen printing the front surface of the cells and the second for printing the rear surface of the cells.

Anti Reflection Coating

To enhance the short circuit current of the solar cells, and hence their power output and efficiency, the reflection from the top surface should be reduced. This can be achieved by having a suitable Anti Reflection Coating (ARC) on the top surface. There are several techniques for applying such a thin ARC e.g. vacuum evaporation, plasma or spray pyrolysis. The most suitable technique for volume production is spray pyrolysis using suitable organo-metallic compounds.

Solar cell testing

The solar cells are then tested and sorted according to their current ranges, using a computer controlled automatic cell tester and sorter.

SPV module

Although solar cells are the basic building blocks of SPV power generating systems, it is the SPV modules

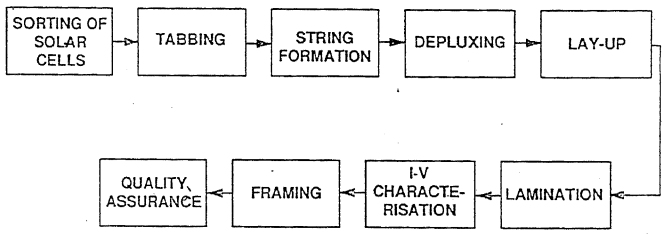


Figure 5. Moduling process flow chart.

and not the solar cell, with which system designers work. The current/power output of a single solar cell is too low to run most electrical appliances. So, several cells have to be interconnected to produce operationally useable power. In addition, the thin solar cells are fragile and so have to be protected from the environment. Thus a suitably interconnected combination of solar cells, when sealed hermetically with a transparent front glass cover, forms an SPV module. These modules can, in turn, again be suitably interconnected in series and/or parallel combinations and mounted on a metal frame structure to form an SPV panel. An aggregate of SPV panels in the field forms an SPV array. In respect of power levels, a typical standard module made of 36 nos. of 4" dia. circular cells will give a power output of 38–42 W whereas, a module made of 36 nos. of 4" pseudosquare cells will give around 45–50 W. The power and voltage levels of panels and array are dependent upon the specific system design and may, in practice, range from tens of watts to several hundred or thousands of watts of power and from a few volts to several hundreds of volts. Electricity generation from the SPV power source is totally modular in nature, and the SPV power capacity and system voltages can be tailored to meet the requirements of the user.

Module manufacturing process

The solar cell moduling process flow chart is shown in Figure 5. The individual process steps are as follows:

Lead attachment and stringing

Leads are attached to each solar cell by a soldering process. The lead attachment is done either manually or through a semi-automatic tabbing machine. After lead attachment, a definite number of cells are interconnected in series to form a string of solar cells in a defined pattern to give a sizeable output power. The stringing or cell interconnecting operation is also performed by a simple soldering process. Automatic/semi-automatic equipment are also available for performing these operations. An automatic tabbing machine in CEL's moduling line is shown in Figure 6.

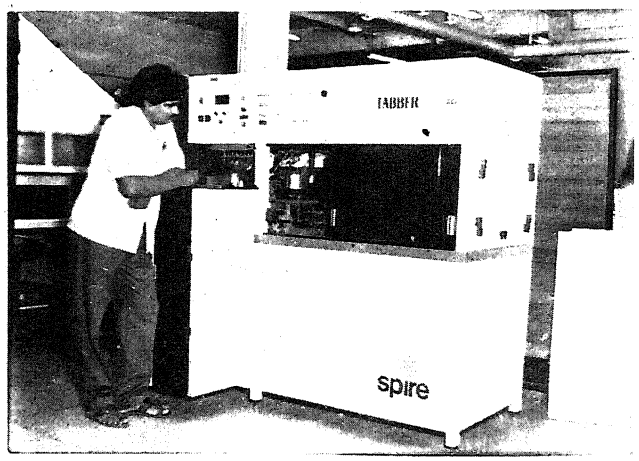


Figure 6. Tabbing machine for soldering leads to cells.

Lay up and encapsulation

The string of solar cells is then chemically cleaned and placed between a sheet of a special soft polymer called ethyl vinyl acetate (EVA) and a sheet of poly vinyl fluoride (PVF) which act as encapsulant substrate for the strings. A sheet of high transmission toughened glass is used as the substrate. The complete lay-up is then put in an encapsulating equipment called a laminator in which it is subjected to a specific time–temperature curing cycle to form the laminated module. While so far the lamination was being carried out using imported machines, indigenous laminators using state-of-the-art features including microprocessor control have now been got fabricated at a much lower price for carrying out this production step in CEL.

I. V. characterization

The laminated module is then tested in a sun simulator in which the module is exposed to a uniform illumination simulating the sun's insolation on a standard midday (i.e. normally designated as 'one sun' i.e. 100 mW/cm^2). Through electrodes attached to the module, the voltage across the module (V) and the current from it (I) are measured at different load conditions from 'open-circuit' (OC) condition to the 'short-circuit' (SC) condition and the complete I–V curve for each module is plotted automatically using a PC. Each such I–V curve is stored in the computer memory after suitably tagging it with the module number and other batch data. I–V characteristics of a typical module are shown in Figure 7.

Framing and termination

The encapsulated module is framed using aluminium channels of special design to protect the module from mechanical damage, and for mounting it on suitable

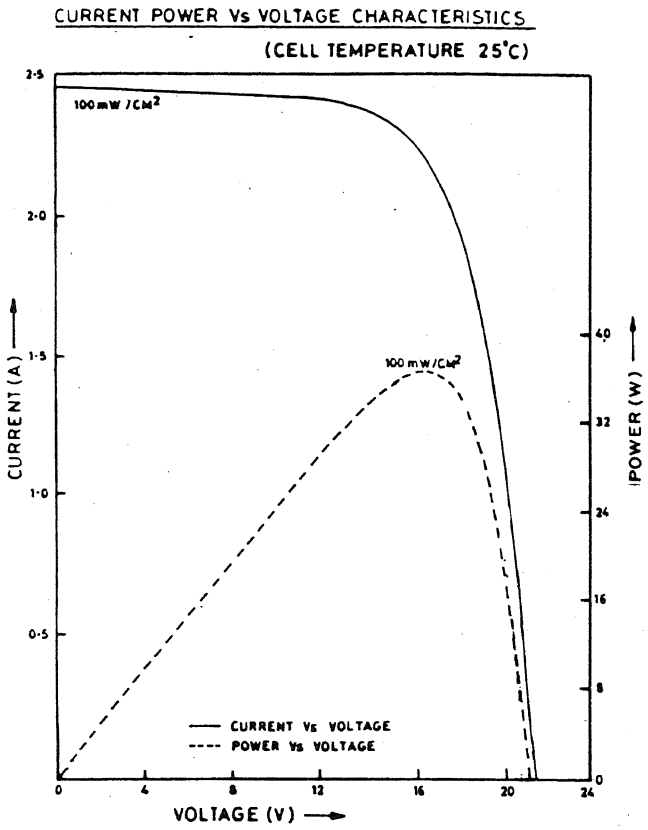


Figure 7. I–V characteristics of a typical module (36 No., 4" circular cells).

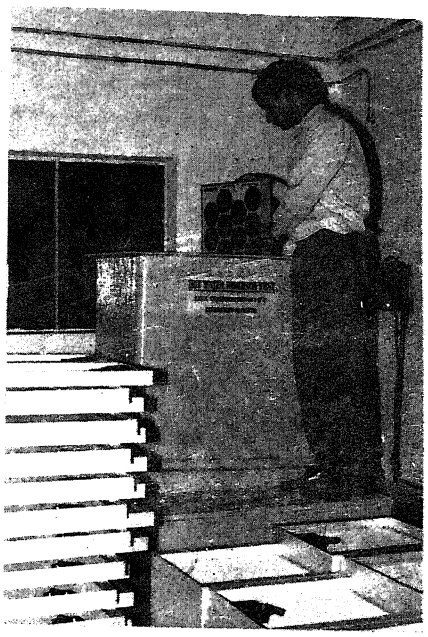
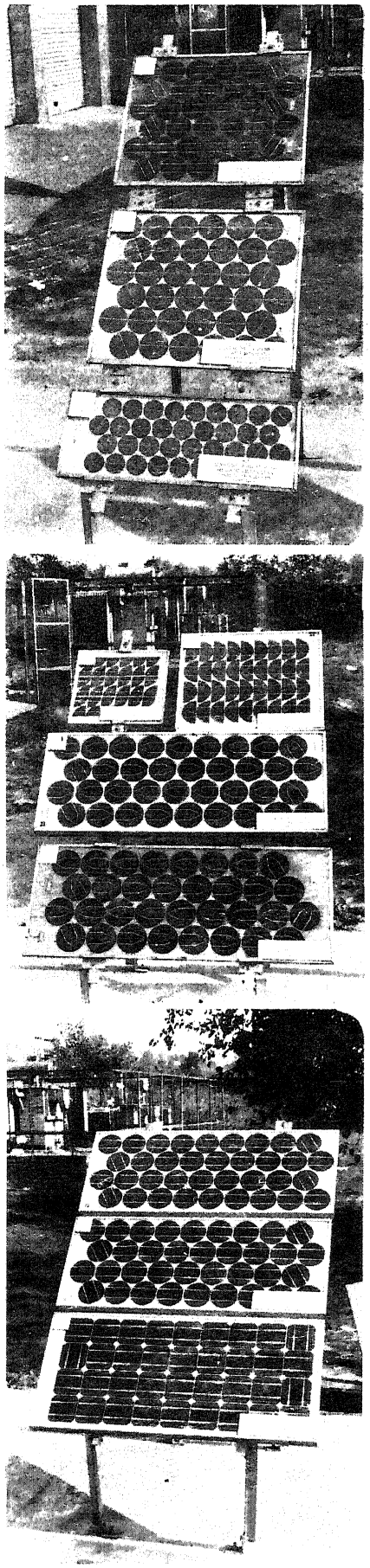


Figure 8. Hot water immersion test.

structures to form a panel. A nylon terminal box is then fixed to provide electrical terminals to extract electrical current from the module.



Quality assurance

A statistically selected number of modules from each production batch are then subjected to a number of quality assurance tests as also environmental tests such as hot water immersion test, hot/cold chamber tests, heat cycle tests, etc. Figure 8 shows the hot water immersion test equipment. The modules are now ready for use as a power generating source either individually or formed into a 'panel' or an 'array' depending upon the specific system design requirements.

The milestones of development of different types of modules keeping pace with the corresponding cell/module process technology milestones are given in Table 2.

Some of the typical modules from this product range are shown in Figures 9–11. The ONGC module is a pioneering instrinsically safe double glass module developed specifically for operation in explosion-prone environments such as on the off-shore, oil production platforms of the Oil and Natural Gas Commission (ONGC). These modules have also been certified to be intrinsically safe in Gr. I, Gr. IIA and Gr. IIB atmosphere hazardous area, by the Central Mining Research Station (CMRS), Dhanbad and accepted by the international insurers, Lloyds of UK. These are the first modules in the world with such certification.

As a result of all this technology development, CEL has presently the following types of modules in its regular production range:

	Type of module	Power output
a)	36 nos. of 4" circular cell module	38–42 W
b)	36 nos. of 4" pseudosquare cell module	45–50 W
c)	36 nos. of half circular cells of 4"dia.	10–14 W (8 V)
d)	36 nos. of half cells of 4" circular/ pseudosquare	10–20 W (16 V)
e)	Lantern modules – using 1/3, 1/2, 1/4 of 4" circular/pseudosquare cells	10 W (8 V/16 V)
f)	Lightweight foldable manpack module	24 W (16 V)
g)	Lightweight 'book case' type foldable module	14 W (14.5 V)
h)	Radio/tape-recorder module – 18 nos. of cut cells	3 W (8 V)

Solar photovoltaic (SPV) systems

The energy generated from the solar cell/module can be usefully harnessed to meet any specific load requirement through a properly configured SPV power system

Figures 9, 10 and 11. Different types of modules in the evolution of the technology.

Table 2. Milestones of development of modules at CEL

First module of 1" diameter solar cells	1977
First module of 2" diameter solar cells	1978
First module of 3" diameter solar cells	1980
Intrinsically safe, double glass, stainless steel frame module for off-shore platforms of ONGC	1982
First module of 4" diameter solar cells	1984
First module using lamination technology	1986
First module of 4" pseudosquare solar cells	1989
High voltage module for solar PV power plants	1992
Foldable, light weight, polymer encapsulant defence modules to operate from − 30°C to + 55°C	1993
Lantern module	1993

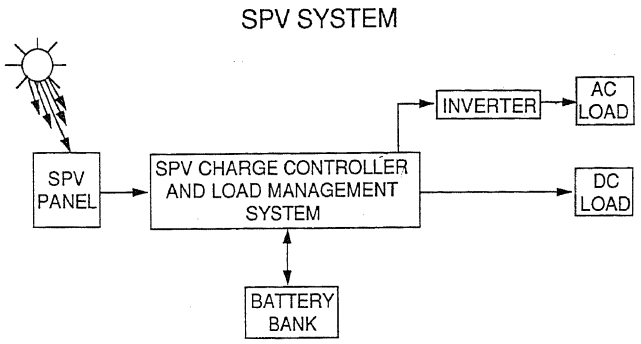


Figure 12. A simple schematic of an SPV system.

consisting of the SPV modules/panels/arrays and other Balance of System (BoS) components such as battery bank, charge controller, inverter and loads such as pumps, lamps, vaccine refrigerators, TV sets, radio communication equipment, Very Low Power TV transmitters etc. A simple schematic of a typical SPV power system is shown in Figure 12. The SPV panel generates DC electricity from the sunlight incident on it. Where the SPV system is to operate during the daytime e.g. a solar water pump, the electricity generated by the panels can be directly used to run the DC pump. On the other hand, where the application is at night, such as powering a light, the electricity generated during daylight hours is stored in a battery through appropriate charge regulating electronic equipment (charge controller) and the power is drawn off the battery at night. The DC power generated can directly be used to run DC load devices. If the load is an AC device, the DC power generated by the panels (or from the battery bank) is converted into AC power of the desired phases (single or three-phase) and frequency (normally 50 Hz), through an inverter and used to power the load/equipment.

SPV systems development at CEL

The design of SPV systems are load and location-specific. SPV power systems are custom-designed for a

Table 3. Milestones of SPV systems development at CEL

SPV operated radio set in Jammu and Kashmir	1978
SPV irrigation water pumping system installed at Kalyani, West Bengal	1979
SPV pumping system for drinking water near Jaipur	1980
SPV community lighting systems installed at SOS children's village Leh, Ladakh, J and K	1980
SPV community television system installed at Kalyani, West Bengal	1980
Special SPV power pack for the first Indian scientific expedition to Antartica	1981
First SPV power system for unmanned offshore oil-well platform of ONGC	1982
SPV operated diesel dispensing pump	1982
Export of SPV water pumping and other systems	1983
First SPV powered stand-alone street lighting systems in Salojipalli, Andhra Pradesh	1983
SPV foldable modules for Everest-84 expedition	1984
SPV systems for direct reception community TV sets of Doordarshan	1985
SPV system for cathodic (anti-corrosion) protection of oil pipelines	1985
SPV system for vaccine refrigeration in village health centres	1985
Design and supply of first SPV system for very low power unmanned TV transmitters (VLPT) and TV Receive only (from Insat) (TVRO) terminals of Doordarshan	1986
Supply of SPV power system for weather monitoring, lighting and entertainment at an avalanche study station at Rohtang pass (alt: 3600 m/12000 ft.) in the Himalayas	1987
SPV powered lighting and television system for army camp at Siachin in the Himalayas (alt: 3600–5500 m/12000–18000 ft.)	1987
Supply of first 2 kWP SPV village power pack	1987
SPV system for warning signals at unmanned railway level crossings	1988
SPV rural deepwell water pumping systems for national drinking water mission	1989
SPV power systems for rural telecommunication transmission and switching equipment	1989
Supply and installation of SPV systems in village Hitam-ulu, Sumatra, Indonesia	1990
SPV lighting systems for adult education under National Literacy Mission	1990
Supply and installation of 6 kWP power plant at Kariawali village in Uttar Pradesh	1990
Supply and installation of Asia's first 100 kW SPV power plant at Kalyanpur village near Aligarh, UP	1992
10 kW power generator for M. S. Swaminathan Foundation at Taramani, Madras	1993
Supply of foldable lightweight SPV modules for powering manpack communication equipment for the defence services to operate from − 30°C to + 55°C	1993

specific application at a particular geographical location, by taking into account both the intensity and diurnal/seasonal variation of solar insolation at that place. The total SPV panel/array capacity (in watts or kilowatts) and battery capacity (in ampere-hours) are accordingly chosen to meet the total load current, duration of daily use and the above mentioned solar insolation data. Automatic switching 'ON' and 'OFF' of the power from the SPV system to the load (say lights during night time) is also incorporated in the system design to restrict the usage of the SPV power to the exact duration of power requirement. This is to save the life of the system,

particularly of the batteries and electronic BoS. Development of appropriate SPV systems to meet different applications and user needs is a primary requirement for wider promotion of the market for SPV power systems. Therefore, parallel to the development of the process technology to manufacture cells and modules, CEL has put in a major effort of development-cum-manufacture-cum-installation and commissioning of a range of SPV energy systems. CEL has continuously been developing a variety of SPV application systems, subjecting them to extensive field testing and then obtaining user acceptance. Thereafter, these systems are produced and supplied against user demands. Indeed the systems constitute the total technology because it is through such systems that solar electricity is actually applied to meet social and economic needs.

CEL has developed a computer software package for designing site and load-specific, stand-alone, SPV systems. The software enables the drawing up of a complete system specification based on site information such as the lattitude and longitude, solar insolation and load parameters including the duration of daily usage as also provision of battery storage for a specified number of 'successive sunless days'. The system design then includes the selection of an appropriate module type (using circular cells or pseudosquare cells for e.g.) for use in the total SPV panels/arrays for the system. For this reason, a close interconnection may be perceived between the system milestones and the module product development milestones.

The milestones of such SPV systems development and introduction into the SPV market by CEL is given in Table 3.

System for ONGC platforms and TV transmitters of Doordarshan

The cases of the SPV systems designed and developed for the off-shore oil production platforms (Figure 13) of the Oil and Natural Gas Commission (ONGC) and for the Very Low Power TV Transmitters (VLPT) of Doordarshan (Figure 14) which are among the unique SPV systems supplied to customers as 'world-firsts' had more than one distinctive feature. The first of those features was CEL's close interaction with the users – ONGC (through Engineers India Limited) and Doordarshan – and with the equipment manufacturers Gujarat Communications and Electronics Ltd. (GCEL), Baroda and Electronics Corporation of India Ltd. (ECIL), Hyderabad for the UHF radio communication equipment and the telemetry equipment respectively for the ONGC Oil Platforms and the GCEL again for the VLPTs for Doordarshan, to match the SPV power system with the electronic equipment concerned. Secondly, through such interactions, the power consumption of the concerned

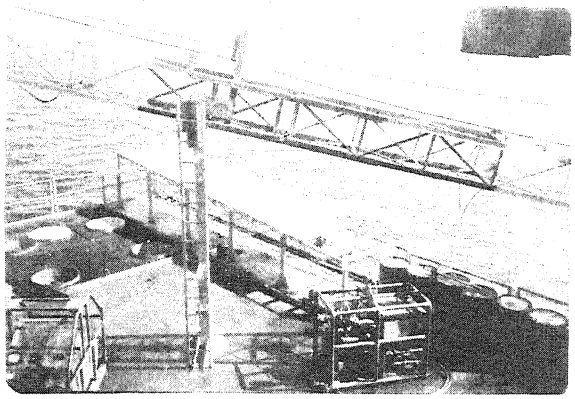


Figure 13. SPV panels on the off-shore oil well head platform of ONGC.

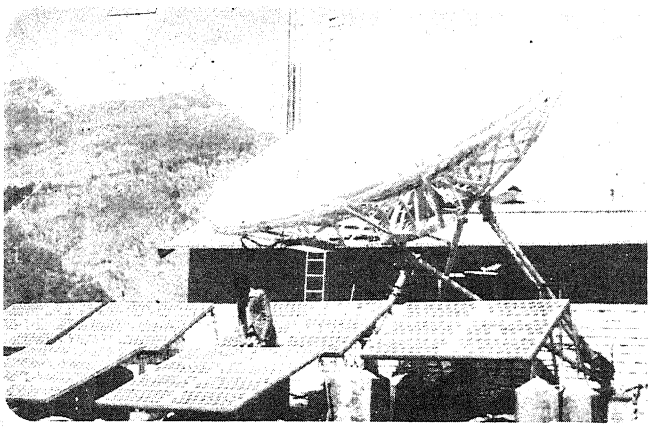


Figure 14. The panels of the SPV power source for the Very Low Power TV Transmitters (VLPT) of Doordarshan.

electronic equipment was reduced (in the case of the VLPT, to almost 50% of the original level) so as to make operation of the equipment with SPV power maximally cost-effective. More recently, a similar exercise is being done with the SPV lantern manufacturers so as to come out with an optimized least-cost product.

Categorization of SPV systems

SPV systems can be categorized on the basis of their application characteristics as rural, remote area and industrial systems. Such a categorized list of SPV systems developed and supplied by CEL is shown in Table 4.

A majority of the SPV systems meet the electricity needs for water pumping, village street and domestic lighting, vaccine refrigeration for immunization and other basic needs of rural areas not covered by conventional grid electricity. With the increasing use and

Table 4. Applications of SPV power systems

Rural applications
Village street lighting
Domestic and community lighting
Shallow well water pumping for irrigation and drinking water supply
Deep well water pumping for drinking water supply
Water purification systems
Village power plants
Vaccine refrigeration
Rural radio telephone sets and telephone exchanges
Very low power unmanned television transmitters
Direct reception television sets
Black and white and colour television receivers
Transistorized radio receivers
Solar lanterns
Remote Applications
Applications in Antarctica
Application in the Mount Everest expedition
Offshore telecommunication and telemetry
Hydro-meteorological stations
Satellite-based data collection platforms for meteorology and oceanography
Industrial applications
Cathodic protection of oil, gas, water and other pipelines
Powering of electronics on lighthouses
Obstruction warning lights at airports
Railway platform lighting and panel interlocking
Warning signal for unmanned railway level crossings
Petrol/diesel dispensing stations
Solar power generator
Defence applications
Manpack battery chargers
Applications by paramilitary forces

awareness of SPV power as an alternative source of energy, the battery manufacturers have come out with long life, maintenance-free batteries, specifically for use in SPV systems. This has rendered the SPV systems much more reliable and longer lasting for use in remote inaccessible areas.

SPV systems for the National Technology Missions

One of the major contributions of CEL's SPV power systems are to the National Technology Missions. Some 205 SPV-powered deep well water pumping systems using submersible pumps have been supplied for the drinking water mission through the Department of Rural Development. These pump systems are really complete village water supply systems and are capable of discharging 10,000 litres/day from depths as great as 120 m. Over 1,500 lighting systems using high luminosity Compact Fluorescent Lamps (CFL) have been supplied to as many Adult Education Centres in 14 states under the National Literacy Mission. Over 35,000 nos. of SPV power sources have been supplied for the Department of Telecommunication's Single Channel VHF Radio Links as part of the National Mission for Rural Communications. Solar-powered refrigerators are also being developed and supplied for the cold chain for storing vaccines and life-saving drugs in remote villages under the National Immunization Mission. Figure 15 shows CEL's SPV-operated lights in an Adult Education Centre at Salem District, T.N.

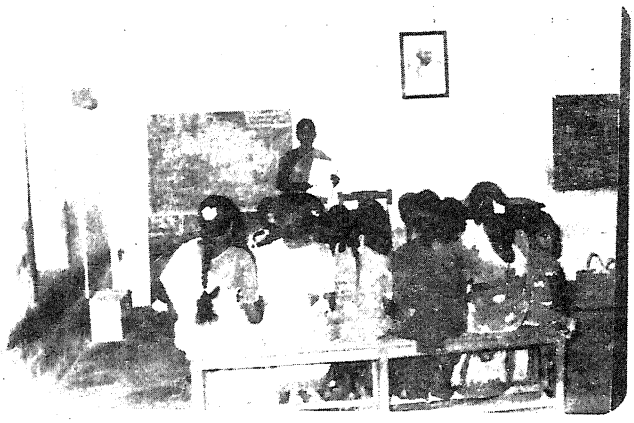


Figure 15. SPV operated lights in an Adult Education Centre at Salem District, Tamil Nadu.

Defence charger

The other pioneering development of CEL has been the 'Defence modules'. These are lightweight, foldable modules which can be carried 'manpack' by the soldier for charging the batteries of communication equipment (wireless radio sets) carried by him. These 'defence modules' have been type-approved by the Army, after exhaustive and extensive trials in actual field conditions ranging from the Rajasthan desert to the extreme cold weather conditions of the high Himalayas (3600–5500 m/12000–18000 ft.). Based on the feedback from such field trials by the customer, CEL has optimized the design, engineering and production technology of these modules. These foldable 'manpack' SPV battery charges have now been commercially manufactured and first supplies made to the Army. This defence-approved foldable module is the only one of its kind in the world. Figure 16 shows CEL's foldable modules in use at a base camp for charging batteries of communication equipment in the Indian Everest Expedition of 1984. The range and total number of SPV systems developed and supplied by CEL so far are shown in Table 5.

CEL has been responsible to a large extent for the creation of a national market for SPV products through its continuous and innovative development of new SPV application systems, even when no possible market for these systems existed.

Table 5. Solar PV systems supplied by CEL as on 31.3.95

1.	Stand alone street lights	11095
2.	Indoor lighting systems	251
3.	Lighting-cum-TV receiver systems	449
4.	Lighting systems for rural adult education centres	1558
5.	Solar PV lantern	1051
6.	Shallow well village water pumping systems	947
7.	Shallow well village water pumping systems (National Pump Programme)	92
8.	Deep well village water pumping systems	205
9.	Multi-use village power packs/micro power stations	21
10.	Vaccine refrigeration systems at rural health centres	30
11.	Terrestrial and direct-from-insat TV receivers	1253
12.	SPV power source for Very Low Power (unmanned) TV transmitters/transposers	211
13.	SPV power source for DOT's multi access rural radio communication systems	1831
14.	SPV power source for DOT's single channel VHF radio telephone equipments	31029
15.	Off-shore oil/gas platforms	27
16.	Cathodic protection systems for oil pipelines (to prevent corrosion)	12
17.	Obstruction warning lights systems for civil aviation	34
18.	Unmanned railway level crossing warning systems	20
19.	SPV systems for powering of compound lights/garden lights and water pumping system	1
20.	SPV system for diesel/petrol dispensing stations	15
21.	Defence systems	309
22.	Others	422
		11095

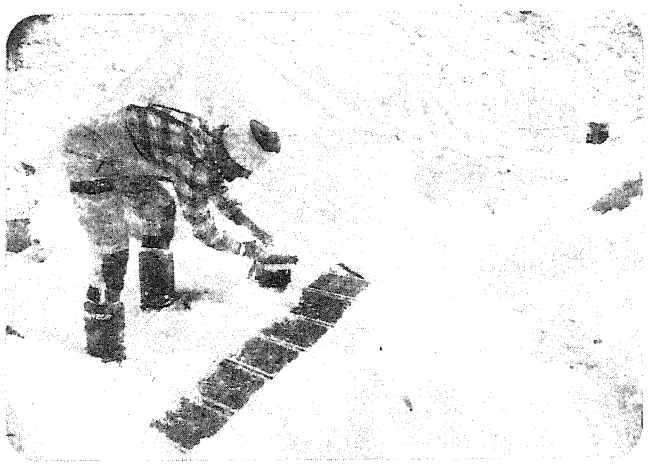


Figure 16. Foldable modules in use at a base camp for charging batteries of communication equipment in the Indian Everest Expedition of 1984.

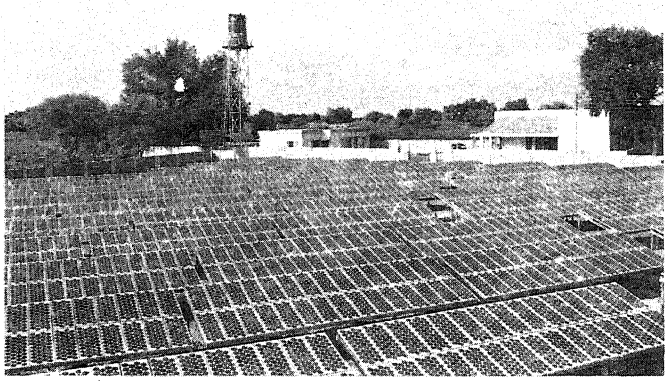


Figure 17. The SPV array of the 100 kW plant at village Kalyanpur, Aligarh District, UP.

With the development and technology and technology-proving phase behind it, the SPV market could now be truly said to have come of age both in our country and world over with ever-increasing use of solar photovoltaic systems and power plants coming into vogue. Such a growing market is expected to lead to further cost-reduction of SPV power as a result of both economies of scale and technological advances.

100 *kW power plants*

CEL has pioneered during 1992–94 the design, engineering, construction and operationalization of two 100 kW SPV power plants, Asia's largest such plants. The plants are located at villages Kalyanpur and Ghosi in the state of Uttar Pradesh. Figure 17 shows the SPV array of the power plant at Kalyanpur. These power plants have been fully designed by CEL and installed on a turnkey basis against contracts from the Non-Conventional Energy Development Agency (NEDA) of the Uttar Pradesh government. The power plant at Kalyanpur powers 450 domestic lights, 50 street lights and 15 nos. of 5 HP water pumps. CEL has set up the plant on a turnkey basis including supply of the special high voltage SPV modules, the lights, pumps, and the Power Conditioning Units (PCU), building the control room to house the PCUs and batteries, and laying the cables and interconnecting the whole system etc. The plant is designed on a modular basis so as to consist of 4 nos. of 25 kW power units including the corresponding SPV sub-array, PCU and batteries. The PCU contains a charge controller for safeguarding the battery life and an inverter for converting the DC electricity from the SPV array/batteries into 3-phase, 440 V, 50 Hz AC. The PCUs also have a phase synchronization unit to enable their interconnection in parallel to meet the total load requirement in the stand-alone mode. One 25 kW sub-system has been interfaced with the national AC grid through a 2.5 KMs, 11 kV feeder line. All the PCUs including the one for the grid interfacing have been totally designed indigenously to cater to the type of voltage and frequency variations prevalent in the Indian grid. Power conditioning units available in the international market do not have the capacity of handling such variations. The 25 kW basic unit has now become a

standard module for such plants to be installed by CEL in future. The modular concept ensures built-in redundancy in the power plant operation so as to meet the load requirement (atleast partially) in the event of failure of any particular PCU. The undertaking of these two SPV power plants, the 1st of their kind in the developing countries, has provided a wealth of real field experience to CEL in this area, thus adding to its already wide range of capability in the SPV power systems.

National Pump Programme

This year, a National Programme for large-scale distribution of SPV water pumps to villages throughout the country has also been launched by the Ministry of Non-Conventional Energy Sources (MNES). This programme has been conceived to give concrete shape to the announcement of the Prime Minister, P. V. Narasimha Rao, of large-scale use of solar photovoltaic (SPV) power for the benefit of populations in unelectrified villages. The programme envisages the sale of these SPV pumps directly to the users under a suitable leasing/soft loan mechanism with instalment payments spread over 10–20 years. This is being attempted to offset the disadvantage of the initial high cost of the solar photovoltaic systems which is acting as a prime deterrent for the large-scale deployment of these systems inspite of its many other advantages. The success of this programme will see the emergence of a very large market for the solar photovoltaic systems in India.

Thrust areas for R&D

In order to progressively increase the share of SPV in meeting global energy demands, intensive research and development work is needed on a continuous basis both to refine and bring down the cost of existing manufacturing methods and to bring into manufacture cells using novel technologies.

These and other measures taken on the technology front along with appropriate policy support from the Government, is expected to steeply reduce the cost of solar cells and modules and hence of the corresponding SPV systems in the country, by the year 2000.

Assessment of CEL's experience in 'Bridging the gap'

When one undertakes an 'Operation Hindsight' of CEL's experience with developing and commercializing SPV technology, several interesting features of 'bridging the gap' between physics and industry come to light.

First, in cell-making, cell-interconnecting and module-making, there is a close interplay between process steps and the equipment with which those steps are implemented in industrial production. Almost every production equipment has to be extensively 'tuned' or optimized in terms of parameter settings and operating conditions, in order to get the best results. Second, selection of an appropriate mix of manual operation and (semi) automation, the choice being made so as to minimize the overall process cost for a plant of 2–3 MW of cells per year. Third, in respect of the choice of the source of supply of the process equipment, a judicious mix of indigenous and imported equipment is made to achieve a cost-effective mix without trading off quality and yield. Some equipment are bought in the world market. This is done where technological sophistication and performance with high throughputs are the dominant consideration. In other cases, indigenous equipment are selected without sacrificing the same technical and throughput factors, but at far lower cost. Fourth, even where equipment are imported, modifications need to be effected in the equipment after use in the plant so as to optimize the equipment into the process line, thereby increasing its effectiveness and performance. Notable examples of such modifications done by CEL are in the tabber and laminator. Considerable modifications were effected in the machine while using it in CEL's line by which its performance increased several-fold. The knowledge acquired by CEL's engineers in making these improvements also provides a sound base for such equipment to be designed and manufactured in our country with substantial cost savings. Fifth, as the process is an integrated one, there is considerable interaction and interactive feedback between one production step and atleast the immediately preceding and succeeding steps in the total process sequence. Continuous iteration is therefore needed, *along the whole process sequence*, to achieve the highest performance solar cells. Sixth, and very crucially, continuous in-process measurements and off-line specialized measurements are needed for both process optimization and control and for product quality control and assurance. The science-base is ever-

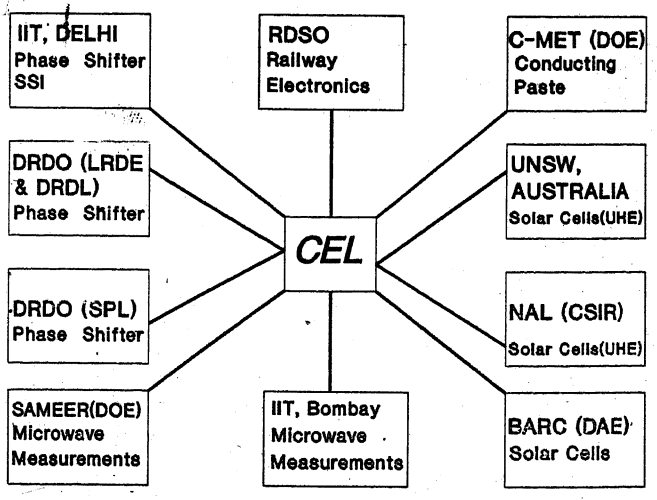


Figure 18. CEL's technology linkages.

present. Seventh, the whole process of technological advance is incremental and evolutionary in character i.e. relatively small improvements are being made all the time to improve process yield and product i.e. solar cell and module quality and performance. This has to be a continuous drive undertaken endlessly. Eighth, as the technology is continuously evolving in this way, the industrial work force operating the plant has to be extensively trained to do two things simultaneously: operate the plant on a disciplined basis with process and machine settings which have been specified as the 'optimum', at a particular point of time. However, to be open, at the same time to apply modified settings as the R & D effort (which is going on all the time) reveals the need for altered process parameters in order to achieve improved performance, whether of efficiency, yield or quality. This calls for flexibility of response and the collection, review of large amounts of data on a continuous basis.

External linkages

In the evolutionary development of technology described in the foregoing, the development process has been largely undertaken in-house by CEL's SPV group. However, there has also been close interaction and help from external agencies e.g. research laboratories, test houses and indigenous and foreign equipment manufacturers. This interaction is schematically shown in Figure 18.

Conclusion

The case of CEL's experience demonstrates the importance of entering an area of advanced technology at an early stage in the evolution of the technology and of the building up of indigenous capacity to convert science into technology and further, technology into industrial production capacity.

It also clearly brings forth the importance of pilot and demonstration phases in proving the new technology and the capability to optimally mix process know-how, equipment engineering and plant engineering.

As a result of sustained work over the last 15 years (1977–92), CEL has built up considerable capability in the SPV area of integrating science, technology and industry, rather than generating the scientific knowledge and then trying to bridge the gap to industrial application.

1. Becquerel, A. E., *C.R. Acad. Sci.*, 1839, **9**, 145–149.
2. Adams, W. G. and Day, R. E., *Proc R. Soc.*, 1877, **A25**, 113.
3. Ohl, R. S., 'Light-sensitive Electric Device' US Patent no. 2, 402, 622; Light-sensitive Device Including Silicon, US Patent no. 2, 443, 552; both filed 27th May, 1941.
4. Kingsbury, E. F. and Ohl, R. S., *Bell System Tech. J.*, 1952, **31**, 802–815.
5. Chapin, D. M., Fuller, C. S. and Pearson, G. L., *J Appl. Phys.*, 1954, **25**, 676–677.
6. Green, M. A., Wenham, S. R., Zhao, J., Zolper, J. and Blakers, A. W., Proceedings of the 21st IEEE Photovoltaic Specialists Conference, Kissimimee, Florida, USA May 21–25, 1990, p.207.
7. Green, M. A., Wenham, S. R. and Zhao, J., Proceedings of the 23rd IEEE Photovoltaic Specialists Conference, Louisville, Kentucky, USA. May 10–14, 1993, p.8.
8. Saitoh, T., Shimokawa, R. and Hayashi, Y., Proceedings of the 4th Photovoltaic Science and Engineering Conference, Sydney Australia, February 14–17, 1989.

Mushrooms: Beauty, diversity, relevance

C. V. Subramanian

The literature on mushrooms is scattered, vast and rich. Current knowledge of mushrooms is essentially a synthesis of mushroom legends and myths, cults and folklore, re-inforced by mushroom biology and science. Mushrooms, indeed, are beautiful: beauty derives not merely from elegance and variety of form and colour, but structure and architecture superbly linked to function and life-style. Of ancient lineage, as decomposers, symbionts, and root-infecting pathogens, mushrooms have evolved with plant life through the ages. They are thus a unique component of the biota. Their diversity is a reflection of their lineage and their evolution. Mushrooms have merit, and meaning: species range from being delicious to edible to hallucinogenic to poisonous, even of potential value and use in medicine. Their enumeration, taxonomy, distribution, biology, chemistry, cultivation and conservation, naturally, are of great relevance and will remain of perennial interest and fascination for us.

THE theme of this paper reflects man's perennial fascination for mushrooms from times immemorial. Mushrooms are of ancient lineage, omnipresent, remarkably beautiful and diverse in their form, in their interaction with other biota and in all that they do. The suddenness and immensity of their presence in the wake of thunder and lightning, and their transience and fleetingness had been known since antiquity imparting a sense of mystery about the mushroom in man: amazed, he holds the thunderbolt, the thunder God responsible! Legend apart, the observation is in place. Impelled by curiosity, he notes that, of the many around him, some may be eaten and, with experience, he soon finds that some can be poisonous and so cannot be eaten. It seems plausible he also knew how to distinguish the edible from the poisonous, an exercise in taxonomy that calls for reliability, absolute reliability, ensuring safety. He does not stop there: the mushroom must be domesticated, like plants, must be grown, cultivated. Indeed, Dioscorides showed the way: the ancient method of sprinkling mushroom spores on to moist wood and waiting for the fruiting bodies to appear is apparently crude, but yet is the forerunner of modern methods of mushroom cultivation; it also marks the beginning of our insight into the biology of the mushroom. Experience with a diversity of mushrooms eventually leads to the recognition, in some, of inebriant qualities which, in time, we see, are built into religious practices or cults in many parts of the world. The use of mushrooms by primitive peoples is now well documented and the rise of the subdisciplines, ethnomycology and ethnopharmacology, and the corpus of data they present on the chemistry, metabolism, and biological activity of hallucinogens and other bioactive compounds is clear evidence of the remarkable intuition and ingenuity of primitive man in discovering the true properties of these mushrooms. Indeed, legend, myth and science all now become part of a totality that heightens the meaning and fascination of mushrooms for man.

One would imagine it was easy for early man to collect mushrooms in the wild to satisfy his needs. The practice of collecting mushrooms in the wild persists to this day in many parts of the world. Man's early concern with the natural occurrence of his favourite mushrooms, the habitats where they occur, the periodicity or the seasonal character of their fruiting, their abundance, and so on, indeed, marks the beginnings of mushroom sociology or ecology.

Mushroom biology

The occurrence of mushrooms on such familiar substrates as wood, litter and soil, implies a role for them in these microhabitats. Forest litter and forest soils are often literally permeated by fungal threads or tubes (collectively, the 'mycelium') often forming 'rhizomorphs', capable of free and extensive spread in litter and in soil. The distribution and development of 'mycelial cord systems' of decomposer mushrooms are often conspicuous in soil and in litter of deciduous woodlands. Many of these are what we may call true inhabitants of the soil though the routine soil dilution plate technique widely used by soil microbiologists to isolate fungi from soils may not reveal their presence in soil. These fungi may obtain food directly from the substrate which then serves as a food base for the mycelium of the fungus spreading radially therefrom.

C. V. Subramanian lives at 'Anjaneya', Plot 885, 62, Ramaswamy Salai, K. K. Nagar, Madras 600 078, India.

They have the enzyme systems required to break down substrates, especially lignin, hemicellulose and cellulose which are the main components of litter and of wood in the forest floor. In fact, degradation of lignin and wood is primarily a function of the great fungal group, the Basidiomycotina, to which all mushrooms belong, besides the pore fungi or polypores ('bracket fungi') and a host of others, more modest only in their general appearance and profile. Collectively, then, they are key scavengers of so much of the organic material in forest floors and soils, of agricultural and other wastes, all of which must be recycled for the sustenance of the biota. In short, the basidiomycetes are a key component in re-cycling of nutrients in many terrestrial ecosystems. Their biomass in forest soils and in forest litter is not easily quantified, but must be considerable. What is more, many mushroom species, and also other specific mycota, are symbionts of tree species, often obligately so, so that the symbionts in this remarkable partnership depend on each other for development and sustenance. That is a further indication of their relevance in forestry and afforestation programmes. Symbiosis leads to the development of mycorrhizas (literally, 'fungus roots') and there are innumerable such mycorrhizal associations ranging from the loosely specific to the not so specific. From the occurrence and characteristic distribution of their fruit bodies one might discern associations with particular tree species, reflecting a possible mycorrhizal relationship. Mycorrhizal fungi mostly are species of genera such as *Amanita, Boletus, Lactarius, Russula* and *Tricholoma*. The well-known fly-agaric, *Amanita muscaria* forms mycorrhizae with *Betula*, *Pinus* and possibly with *Carpinus*. Production of fruit bodies by mycorrhizas depends upon supplies of host-assimilates, and on host genotype, age of the host tree, soil type and soil amendments, and on weather and climate. The diversity of mycorrhizal fungi on a given tree species may increase with ageing of the tree species, so that there would be a succession of mycorrhizal associates on a given tree, a reflection of the low specificity of the mushrooms involved in such succession. Some mushroom species such as *Laccaria laccata*, and species of *Boletus* and *Tricholoma* present a dual behaviour, being mycorrhizal symbionts or being saprotrophs capable of decomposing leaf litter.

There is a succession also on dying and dead wood and other substrates. Some species in some genera preferentially colonize specific substrates: many wood-rotters belong to the group, the Aphyllophorales. Those involved in the decomposition of leaf litter and non-woody litter largely belong to the genera *Collybia*, *Clitocybe*, *Marasmius* and *Mycena*.

The mycelium of some species in litter and wood shows bioluminiscence. In fact, the fruit bodies also are luminiscent in some species, e.g. jack-o-lantern (*Clitocybe illudens*), the poisonous moonlight mushroom (*Lamptero-myces japonicus*). In the case of *Mycena rorida* v. *lamprospora* the spores are luminous.

There are also mushroom species such as *Armillaria mellea* that infect living trees, especially root systems, and cause root disease. Root infection may also be due to polypores such as *Fomes annosus*. When the tree has been weakened, colonization by other mushroom species or by polypores and other basidiomycetes may follow. The mushroom group, the Marasmieae, especially the Marasmiineae, comprising species of genera such as *Mycena*, *Crinipellis*, etc. are common on tropical wood and litter, but some are pathogens. *Crinipellis perniciosa* is pernicious, being parasitic on cocoa trees (*Theobroma cacao*) and on *T. bicolor* and the wild species, *T. speciosum*: the fruit bodies are formed only on witches broom deformations of the diseased host. Originally known only in S. America, the disease now occurs wherever cocoa is grown, from Ecuador to the West Indies. *Crinipellis perniciosa* is unique in having a distinct parasitic and a saprophytic phase. The parasitic mycelium in the tissues of the brooms is intercellular without clamp connections; the saprotrophic mycelium arises from the parasitic, following death of broom tissues and has clamp connections: it can then be isolated and grown on media. *Marasmius viegasii* causes a root rot of coffee in Brazil: rhizomorphs on roots of affected trees are conspicuous. Curiously, some mushroom species that are parasitic on trees such as, for example, *Armillaria mellea*, are also symbionts of large achlorophyllous orchids. Some achlorophyllous 'saprophytic' angiosperms, including the Orchidaceae, obtain their carbon compounds through fungi that are parasitic or mycorrhizal on green plants.

The life of the mushroom

The life of the mushroom from the germination of the spore through development of a hyphal system to the production of the characteristic fruit body is a continuum punctuated by events of import. First, nutrients and water are transported from the vegetative mycelium which draws its nutrition from the substrate or from a host. The elongation of the stipe is a result of divisions as well as elongation of individual cells. The fruit body itself is short-lived, and is organized for the production and dissemination of spores. The variations in colour, consistency, configuration, construction and dimensions of the fruit body, and the range in shape, colour and size of spores and other features all contribute to the diversity of mushrooms we see around us. The fruit body size of the most common mushrooms falls within a normal range, but some may be unusually small and others unusually large. The largest fruit body of a mushroom is perhaps that of *Boletus colossus* from the Malagasi Republic: it has a stipe 30 cm tall and 22 cm thick, and a pileus 60 cm wide and 4–6 cm thick, and weighs about 6 kg. Some bracket-fungi are many times larger and weigh much more.

Highly synchronized basidia are typical of the agarics and the polypores; in the latter the hymenium is protected

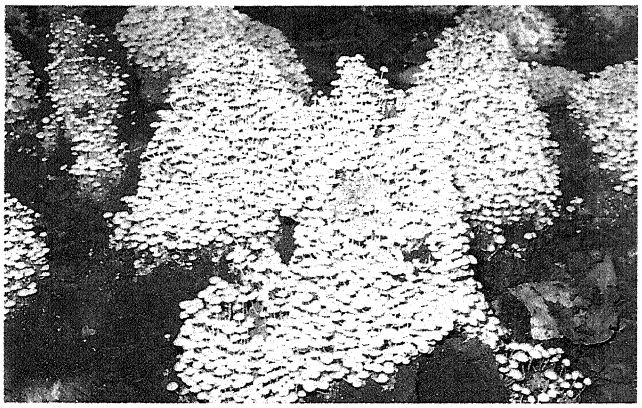


Figure 1. *Termitomyces microcarpus* 'mushrooms' like this.

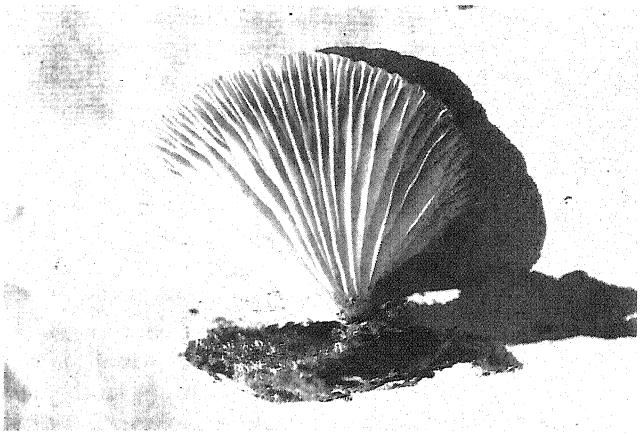


Figure 3. *Pleurotus cystidiosus. Pleurotus* spp. are edible and are cultivated. Note gill architecture.



Figure 2. Polypore on wood. Polypores are great decomposers of wood and play a major role in recycling of nutrients in the forest floor.



Figure 4. *Auricularia polytricha*, the Chinese wood-ear (*mu erh*) fungus. Edible, cultivated. Hong Kong is the principal exporter of this species.

by elaborate trimitic fruit bodies adapted to spore discharge under adverse conditions; in the former well-organized hymenia are formed on relatively fragile fruit bodies that can expand rapidly to discharge the spores during favourable environmental time-frames. The jelly-fungi (Tremellales) are unique in having relatively primitive hymenia, and in their ability to remain viable through time-frames of repeated wetting and drying.

There have been many penetrating studies of the development of the fruit body itself and its morphogenesis, the relevance of which must be obvious to anyone interested in the induction of fruit bodies and increasing their production at will experimentally. Once dikaryotization is accomplished, the dikaryotic mycelium is virtually perennial and persistent and often long-lived, and can produce fruit bodies year after year, season after season. Radial spread and fruiting often produce 'fairy rings', especially noticeable in the case of *Marasmius oreades* and *Chlorophyllum molybdites*. Normally, dikaryotization is a pre-requisite for fruiting, but monokaryotic mycelia may produce fruit bodies under special conditions. Morphogenesis begins with initiation and, following a pause, is consummated by sudden and rapid growth, expansion and maturation. The fruit body appears all on a sudden, and vanishes quickly too. Size and form of fruit body are closely related: smaller fruit bodies usually have slender stipes and larger ones stout stipes. The stipe serves to hold the pileus in position for release and dissemination of spores; it is held erect due to the turgor pressure of its cells. The positive role the pileus plays in the elongation



Figure 5. *Termitomyces heimii.* Tropical, associated with termite nests. Edible; collected in the wild and eaten.

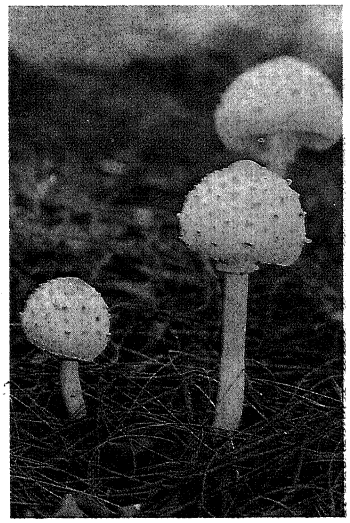


Figure 7. *Chlorophyllum molybdites.* Poisonous. Often forms fairy rings in lawns.



Figure 6. *Termitomyces.* Associated with termite nests in the tropics. The stipe arises from subterranean termite nest. Edible.



Figure 8. *Gymnopillus spectabilis.*

of the stipe first postulated by Schmitz in 1842, is now recognized and has been demonstrated experimentally, for example, in *Agaricus bisporus*, *Coprinus radiatus* and *Flammulina velutipes*: growth control resides in the pileus, specifically in the lamellae.

The way in which the fruit body as a whole, and the configuration and arrangement of the gills and the basidia in the pileus, are functionally adapted to life in a habitat is remarkable. There is a regularity in gill development and gill spacing. The architecture is such that a basidiospore discharged violently falls freely without striking the adjacent gill.

How is this accomplished? As described by the great Buller, to whom we owe much of our knowledge of the biology of the mushroom, the spore trajectory is a spo-

rabola: the horizontal flight is completed in a fraction of a second and is followed by a vertical fall that is gravitational and slow. The system of gills is such that, while they converge from the circumference of the pileus towards the stipe, they all cannot be of the same length as that would make them far apart towards the periphery, if they are satisfactorily apart near the stipe. The solution to this situation is seen in the insertion of shorter gills between longer ones. A spore fall rate of the order of about half-a-million a minute is common for a normal-sized mushroom. Naturally, a spore cloud is the result. By removing the cap from the stipe and placing the cap naturally oriented on, say a sheet of glass, one can have a

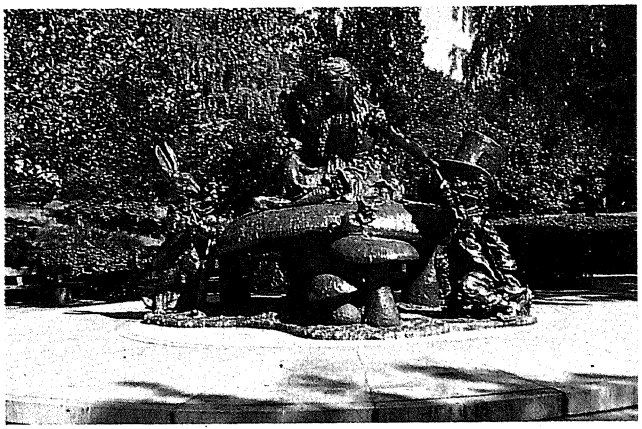


Figure 9. The Monument to Lewis Carrol in Central Park, New York. The sequelae to nibbling the mushroom by Alice, described by Lewis, are well known. Lewis got the idea from the reference to hallucinogenic mushrooms in M. C. Cooke's text-book on fungi.

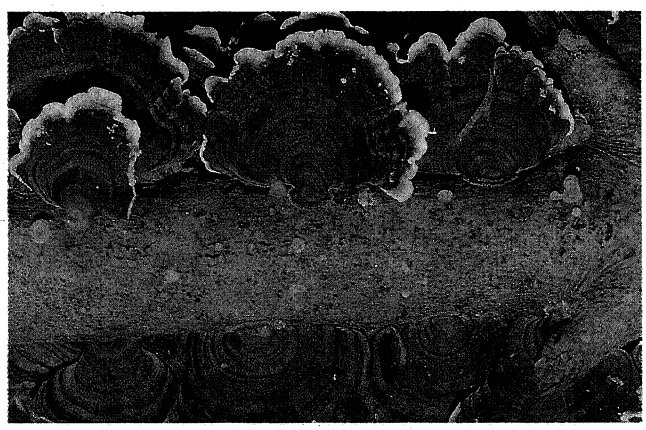


Figure 10. *Coriolus versicolor* is highly valued in Chinese medicine. Causes white rot in wood; sometimes a weak parasite, it forms mycorrhizas with *Galeola foliata*, an achlorophyllous liane.

spore print on the glass in a matter of a few hours. The colour of spores is diagnostic.

Edible mushrooms: Occurrence, diversity

That many mushrooms are edible, and some are delicious, has been known, from ancient times. There are Roman mosaics depicting mushrooms. There is a range or diversity of mushrooms in the wild that are edible and, in countries where they are popular, where there are mycophiles, for example, in eastern Europe, collecting in the wild is common. The common wild species include *Can-*

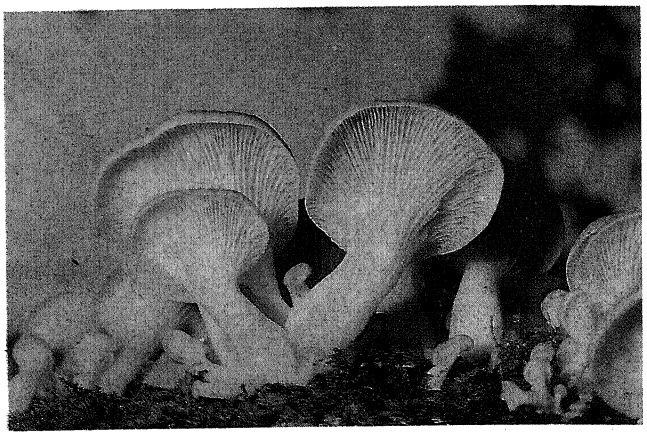


Figure 11. *Pleurotus ostreatus.* Edible, cultivated.



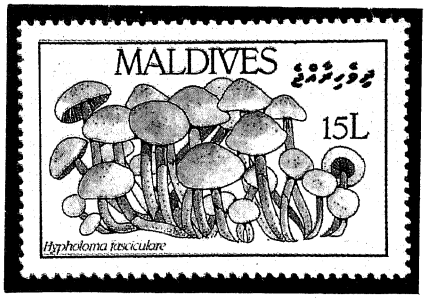
Figure 12. *Ganoderma lucidum* (ling-chi in Chinese, *linjue* in Thai; also known as the Reishi mushroom). Cultivated extensively in China, Japan, Korea, Thailand. The World Reishi Mushroom Institute in New York and the Arunyik Mushroom Centre in Thailand are engaged in standardizing the quality of the mushroom and its medicinal properties and use.

therellus cinnabarinus (Chanterells), *C. cibarius*, the boletes, *Boletus edulis*, *Suillus luteus*, *S. granulatus*, *Leccinium aurantiacum*, *L. testaceoscabrum* • and the Amanitas such as *A. caesarea*. Many of these are not easily cultivated since they are mycorrhizal and fruit bodies are normally produced only in mycorrhizal association. It is then quite reasonable to speculate on the possibility of having, in afforestation programmes, tree species with mycorrhizal associates from among mushroom species that are edible. The introduction into France and Central Europe of the White Pine (*Pinus strobus*) from the United States in the middle of the last century is an example. The new French pine woods showed an abundance of fruit bodies of the delicious *Suillus placidus*: one cannot be certain if the fungus came with the pine from the US or

Mushrooms on postage stamps



















was already present in French soil and now had the host of its choice to colonize. The introduction of the white pine, a blessing in this way, was otherwise, for quite another reason: its high susceptibility to white pine blister rust. However, one can cite other examples of tree introductions with a high measure of success in providing the tree with an effective or suitable symbiont and, in the process, establishing a mycorrhizal relationship productive of a crop of fruit bodies of a prized mushroom. Success in such a venture implies judicious choice of a mushroom that is edible, and knowledge of its taxonomy and biology, especially its behaviour in symbiosis and in fruiting. The diversity from which a choice can be made includes many species such as *Tricholoma matsutakae*, *Amanita caesarea*, *Suillus granulatus*, *Boletus edulis*, *Leccinium aurantiacum*, *Lactarius deliciosus*, *Cantharellus cibarius* and *C. cinnabarinus*.

There are also many mushrooms that are primarily saprophytic and so can be grown in culture and cultivated to yield fruit bodies that can then be harvested. The button mushroom, *Agaricus bisporus* is the most popular, accounting for slightly over one half of the total of world mushroom production. 'Speciality mushrooms' such as *Shiitake* (*Lentinus edodes*), paddy straw mushroom (*Volvariella volvacea*), oyster mushroom (*Pleurotus*, especially *P. ostreatus*: also *P. eryngii*, *P. floridanus*, *P. sojor-caju*, *P. cystidiosus*) and enokitake (*Flammulina velutipes*) are becoming increasingly popular. Shiitake which is the mushroom of choice in Japan, Korea and China, has a long tradition of cultivation: it is cultivated on wood, virtually on logs of wood cut from hardwood trees such as oak, chestnut, etc. – species of *Quercus*, *Castanea*, *Carpinus*, *Alnus*, *Acer*. *Shiia* stands for *Pasania* and the name Shiitake stems from the fact of its frequent occurrence on shiia and the fact that it can be cultivated on shiia wood. Shiitake occurs naturally as a saprophyte on wood of deciduous trees such as oak, chestnut and beech and on shiia in eastern Asia from China and Japan to Vietnam and Cambodia. *Volvariella* and *Flammulina* are sold fresh. The button mushroom and shiitake are sold both in the fresh and the dried forms.

Auricularia polytricha, the ear fungus, and *A. fusco-succinea* – not strictly mushrooms, but both basidiomycetes – are quite popular in many parts of southeast Asia (China, Japan, Thailand, Taiwan and some islands in the Pacific): they are cultivated on logs of wood or sawdust. There are also several species, not yet intensively cultivated but with potential such as *Grifola frondosa*, *Hypsizigus marmoreus*, *Hericium erinaceum*, *Oudemansiella mucida*, and several species of *Termitomyces*. *Termitomyces*, as the very name suggests, is associated with termite nests in the tropics and is collected in the wild and eaten.

The nutritive value of some of the quality mushrooms very nearly equals that of milk. Mushrooms contain high levels of amino acids such as lysine and leucine, carbohydrates, vitamins, and unsaturated fatty acids, with a mineral content exceeding that in fish and meat, and twice that in most vegetables.

Recent studies culminating in the discovery of thermostable components, cytotoxic substances, and antibiotics in mushrooms (see Subramanian, *Curr. Sci.*, 1991, **60**, 290 for Steglich's report of antibiotics in mushrooms) reinforces the need for caution in using raw mushrooms in the dietary, in salads, for example. The occurrence of toxic metabolites such as the thermolabile audiotoxic *volvatoxin* in the paddy straw mushroom and *flammutoxin* in enokitake call for a reassessment of the wisdom of their use as food. Hydrocyanic acid is produced in quantity by some species (e.g. *Marasmius oreades*, the fairy fungus), but from its volatility little of it would remain after cooking.

The genus *Agaricus* to which the common edible button mushroom (*A. bisporus*) belongs, comprises two groups of species, viz. (a) those turning somewhat yellow on bruising (Flavescentes) and (b) those staining red (Rubescentes). The yellow staining species such as the edible *A. arvensis* and *A. bisporus* are known to accumulate considerable amounts of the heavy metals, cadmium and mercury. Other heavy metals known to accumulate in mushrooms include lead, caesium, iron, cobalt, copper, manganese, molybdenum, nickel, selenium, silver, strontium, thallium, uranium and zinc. Accumulation of cadmium and selenium and other heavy metals in shiitake would render it unsafe for human consumption.

The occurrence of the hydrazine derivative *agaritine*, and the diazonium derivative from it, in *Agaricus*, and the fact that the *p*-hydroxymethyl-phenyldiazonium ion and decomposition products of agaritine can be carcinogenic pinpoint the need for caution in the use of these mushrooms. The agaritine content of fresh and processed cultivated mushroom has been studied: the average agaritine content is around 0.088% based on fresh weight. We certainly need more data on the agaritine content in mushroom species and their possible biological effects.

The corpus of data on methods of cultivation, strain improvement, and biotechnology of mushroom production is enormous and cannot even be touched on here.

Poisonous mushrooms: Syndromes, toxins

Poisoning from eating mushrooms has been known since antiquity. As one would expect, there is a variety of syndromes, and a variety of mushrooms involved.

The most deadly and the most notorious is the Death Cap, *Amanita phalloides*, presumably also the mushroom ingeniously used by Agrippina to do away with her husband, the Roman emperor, Claudius Caesar, in her bid to ensure succession to the throne by her son, Nero – ingenious in the sense that the essential properties of *A. phalloides* had been known to the Romans in advance of chemical studies of the toxins which came centuries later. Claudius, in fact, was poisoned with a dish of a mixture of

the edible *A. caesarea* (so named because Caesar was very fond of this mushroom!) and the deadly *A. phalloides* so that the poisoning itself went unsuspected. The Romans, then, knew the incubation period required by the poison to act, as arousal of suspicion had necessarily to be avoided in administering the poison to the Emperor. Knowledge, then, of the poisonous nature of mushrooms and of the precise symptoms of poisoning naturally preceded critical chemical data on the toxins and the *modus operandi* of the toxins. As far back as 1759, the great German naturalist Jacob Christian Schaeffer (of Regensburg: see Subramanian, *Curr. Sci.*, 1991, **66**, 289) underlined the need for chemical studies of the fungi. In Germany, fungal chemistry flourished under the leadership of Zopf, Zellner and Kogl. And more recently, knowledge of the chemistry of the metabolites and toxins of poisonous mushrooms has been strengthened by the work of several scientists, notably Wieland and Faulstich.

The most potent toxic substances in mushrooms are the cyclic tryptophan-containing oligopeptides that characteristically occur in species of the genus *Amanita*. The genus *Amanita* is unique in having species that are edible, some that are poisonous, and some that are hallucinogenic. The 'Phalloides syndrome' refers to poisoning by *A. phalloides*: it is characterized by a latent period of about 10 h (so that the victim remains unaware of the poisoning and, by the time he becomes aware, it may be too late to seek a remedy!) and a two-phase sequel in which gastrointestinal disturbances (severe abdominal pain and colic, nausea, diarrhoea, vomiting) are followed by progressive liver dysfunction. Accelerated pulse, dehydration, cramps in the legs, and a lowering of blood pressure may accompany the primary symptoms. Two groups of bicyclic peptide toxins have been identified: (a) the *amatoxins* that are octapeptides with an indole (R)-sulphoxide bridge, and (b) the *phallotoxins*, that are heptapeptides with an indole-thio-ether bridge. Nine amatoxins and seven phallotoxins have so far been identified. Compared to the amatoxins, the toxicity of phallotoxins to liver cells is quite low. It is now generally agreed that phallotoxins are not involved in the syndrome; on the other hand, most of the effects are ascribed to amatoxins. The damage to the liver is primarily on nuclei of liver cells, specifically, inhibition of RNA-polymerase B. Amatoxins are known to inhibit RNA-polymerase B in all eucaryotes except amatoxin-accumulating fungi and a few *Drosophila* species, some of which inhabit these fungi. Because of their specific inhibition of RNA polymerase B, amatoxins have become an important tool in biochemical studies. About 5–7 mg of amatoxin would be a lethal dose, depending on the body weight of the victim. Ingestion of the equivalent of 50 g of fresh weight of the fungus or a single fruit body, then, can be fatal.

Amatoxins occur in *Amanita verna* and *A. virosa*, besides *A. phalloides*, all of which are myeorrhizal. Amatoxins are also known to occur in several species of the genus *Galerina* which are deadly poisonous: *G. marginata*, *G. autumnaris*, *G. badipes*, *G. beinrothii*, *G. unicolor*, *G. venenata* and *G. sulciceps*. *G. marginata* is common on deciduous wood, particularly on conifers. *G. sulciceps* is tropical and fatal cases of poisoning from this fungus were reported from Java in Indonesia many years ago. Poisoning from species of *Lepiota* several of which contain amatoxins and are deadly poisonous (e.g. *L. helveola*) represents another example of the Phalloides syndrome.

Protein conjugates of amatoxins are ten times as toxic as natural amanitins and yet, the fact that these can be used for raising amatoxin-specific antibodies would now facilitate use of polyclonal antibodies for amanitin immunoassays. A monoclonal antibody, then, could be a potential antidote in the immunotherapy of human poisoning by the deadly Amanitas.

A syndrome with a long latent period and a two-phase course similar to the Phalloides syndrome is often due, not to *Amanita*, nor even a true mushroom, but a species of a fungus belonging to the Ascomycotina. *Gyromitra esculenta* has long been considered edible, as the specific epithet itself suggests. But there have been many cases of poisoning in some parts of Europe: the demonstration of the occurrence of a toxic substance, *gyromitrin* with high toxicity to the liver, confirms that the species is poisonous. On the other hand, gyromitrin is not stable, and most poisonings are apparently from eating of raw or insufficiently cooked fungus. In pure culture, even the mycelium has been shown to have the toxin. It is also likely that there are strains of the fungus that are not poisonous in some geographical regions in Europe, which perhaps serves to explain the conflicting reports about the edibility or otherwise of this fungus.

Another syndrome which has an extended and long latent period and in which impairment of kidney function is the primary damage is from poisoning by species of the genus *Cortinarius*. *Cortinarius orellanus* is a European species and poisoning has been reported from several countries in Europe. The key poison has been identified as *orellanin* for which a bipyridial structure has been proposed; it has also been synthesized. It is a typical nephrotoxin. In acute cases, kidney damage has even precipitated the need for kidney transplants.

A syndrome with intense bouts of perspiration coupled with salivation and lachrymation is characteristic of *Inocybe* poisoning. These effects are ascribed to *muscarine* which has been detected in several species of *Inocybe*, e.g. *I. patouillardii*. Small amounts of muscarine are known to occur in a number of mushrooms, but normally, very large quantities of these would have to be ingested for them to be able to cause any poisoning.

There are then mushrooms that are not poisonous and are edible on their own, but known to produce toxic effects when alcohol is taken after eating them. The common ink cap, *Coprinus atramentarius* is a good example. A sense of being hot and a metallic taste in the mouth accompany a reddening of the face, neck and chest: there may also be a tingling sensation in the legs and the arms,

and palpitation. The active component, *coprine*, has been isolated. Coprine is an inhibitor of the enzyme, acetaldehyde-dehydrogenase in the liver: it inhibits oxidation of the acetaldehyde formed during the metabolism of alcohol to acetate. The resulting heightened content of the aldehyde in the blood shows up in the symptoms. The normal yield of coprine is about 160 mg kg^{-1} of fresh fungus. *Clitocybe clavipes* is another mushroom species that is incompatible with alcohol.

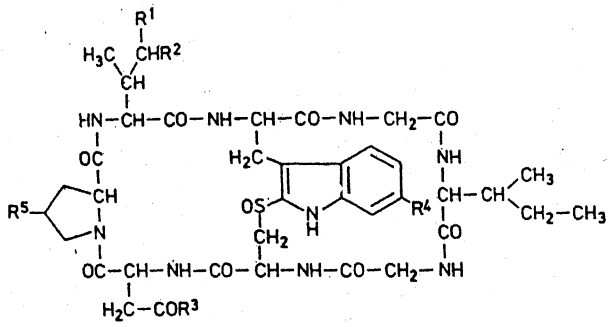
The most widely occurring syndrome from mushroom poisoning is the gastro-intestinal syndrome. According to one report 40% of all notified poisonings in Switzerland during a 40-year period were from gastro-intestinal cases. It should be noted, however, that the syndrome is heterogeneous, being due not to a single toxic species. In general, irritation of the intestinal and gastric mucosae is typical. Several species are suspect in this category: *Chlorophyllum molybdites*, *Hypholoma fasciculare*, *Megacollybia platyphylla*, *Tricholoma pardinum*, and species of *Hebeloma*, *Hygrocybe*, *Leucocoprinus*, and *Macrolepiota*. Toxic principles have been identified or isolated in some species. Thus, *Dermocybe sanguinea* contains anthroquinones such as *dermocybin* and *dermorubin* which are considered responsible. The sesquiterpene, *illudin S.* (*alunamycin*, *lampterol*) is one of the toxic principles isolated from *Omphalotus olearius*, *O. subilludens* and *Lampteromyces japonicus*. The pungency and bitter taste of species of *Lactarius* (Milk Caps) and of *Russula* (Brittle Gills) are well known. Pungency derives from sesquiterpenes, *stearyvelutinal* or *necatorone*. Some at least of these species may be rendered edible by cooking, but those tasting bitter and pungent (e.g. *Lactarius piperatus*, *Russula emetica*) are best avoided. And yet, we must note that in many parts of eastern Europe *Lactarius piperatus* is considered edible.

Poisoning is often due to ingestion of mushrooms, fresh and raw. These include, especially, the following: *Amanita rubescens*, *A spissa*, *A. strobiliformis*, *A. vaginata*, *Armillaria mellea*, *Boletus erythropus*, *B. luridus*, *Entoloma clypeatum*, *Lepiota nebularis*, *Lepista nuda*, *Russula olivacea*. These need to be cooked before being eaten. Edibility even after cooking is questionable for some species such as *Paxillus involutus* and *Sarcosphaera crassa*.

Hallucinogenic mushrooms

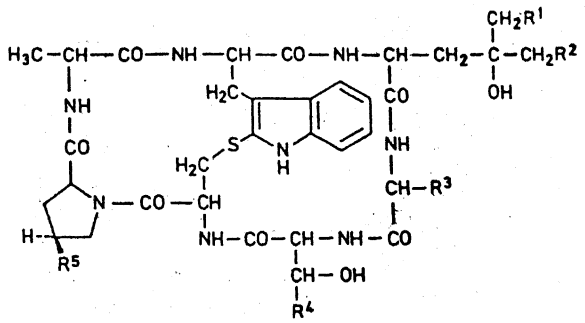
The fly-agaric

The fly-agaric – *Amanita muscaria* – is a mushroom that has perennially arrested our attention by its beauty and its unmistakable identity. It is quite natural that it is popularly illustrated in many a children's book. The fly-agaric is so called from its long-known property of repelling flies, a property now known to be linked to a chemical, the isoxazole, ibotenic acid, first isolated from the mushroom by Japanese scientists. The other properties of this mushroom, quite remarkable, in producing hallucinations on ingestion, have been known for a long time, for example, to the Koryaks and the Chuckchis in Siberia. Their use of this mushroom and the characteristic hallucinogenic effects which are part of the Siberian experience are remarkably well documented by travellers, historians and anthropologists: the accuracy of these accounts is matched by current knowledge of the chemistry of hallucinogenic substances in the mushroom, their metabolism in the human, and their effects on the human nervous system. Wasson reviewed knowledge of the fly-agaric with precision and elegance and even forcefully argued the case for his belief that the fly-agaric is indeed the 'soma' of the Vedas. The first bit of chemistry of the fly-agaric came from the pioneering work of Schmiedeberg and Koppe who isolated from it muscarine in 1869. Concentrations of muscarine in the fungus are, however, low and so, for muscarine poisoning to take effect, one would need to ingest unusually large quantities of fruit bodies of the mushroom. Ingestion of the fly-agaric, nevertheless, produces a characteristic syndrome in which the primary effects are neurological. The effects have been well described by many, notably in the documentation of the use and effects of the fly-agaric by the Koryaks and the Chukchis in Siberia. From its unmistakable features, wide distribution and abundance – being mycorrhizal, one may find it under pines, beech, and birch – it is one of the most widely used psychotropic mushrooms in many parts of the world, but especially in the USA. Symptoms include disturbances of vision, motor excitation, confusion that may be reflected in laughter, dancing, and singing, alteration in perception of personality, and of space and time, of colour, hallucinations, a feeling of illusion and even euphoria. Beginning within about 30 min following ingestion, all is over in about three hours. In the Siberian experience, the Koryaks in their wisdom dry the fly-agaric and it is in the dried form that the mushroom seems most effective. The active principles that are responsible for the effects are now known: apart from ibotenic acid, already mentioned, two other derivatives, muscimol and muscazone, have been discovered. Only ibotenic acid occurs as such in the fungus; the other two are secondary products that appear, for example, quite interestingly, when the fungus is dried. We must therefore assume that the heightened hallucinogenic properties of the mushroom in its dried form must be due to these secondary products. The predominant effect is inhibition of motor functions: muscimol is five to ten times more potent than ibotenic acid. Muscimol is stereochemically γ-aminobutyric acid (GABA) and so its effect on the nervous system stems from stimulation of GABA-receptors of the cells of the central nervous system. A further point of interest is, besides ibotenic acid, muscimol too has insecticidal properties.



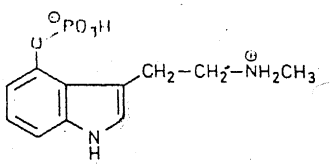
Amatoxins

	R^1	R^2	R^3	R^4	R^5
α-Amanitin	CH_2OH	OH	NH_2	OH	OH
β-Amanitin	CH_2OH	OH	OH	OH	OH
γ-Amanitin	CH_3	OH	NH_2	OH	OH
ε-Amanitin	CH_3	OH	OH	OH	OH
Amanin	CH_2OH	OH	OH	H	OH
Amaninamide	CH_2OH	OH	NH_2	H	OH
Amanullin	CH_3	H	NH_2	OH	OH
Amanullinic acid	CH_3	H	OH	OH	OH
Proamanullin	CH_3	H	NH_2	OH	H

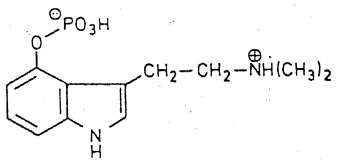


Phallotoxins

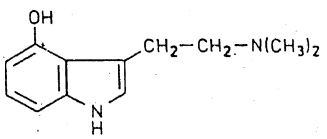
	R^1	R^2	R^3	R^4	R^5
Phalloidin	OH	H	CH_3	CH_3	OH
Phalloin	H	H	CH_3	CH_3	OH
Prophalloin	H	H	CH_3	CH_3	H
Phallisin	OH	OH	CH_3	CH_3	OH
Phallacin	H	H	$CH(CH_3)_2$	COOH	OH
Phallacidin	OH	H	$CH(CH_3)_2$	COOH	OH
Phallisacin	OH	OH	$CH(CH_3)_2$	COOH	OH



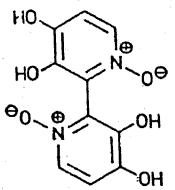
Baeocystin



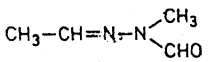
Psilocybin



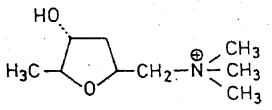
Psilocin



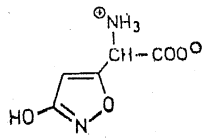
Orellanine



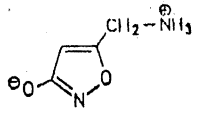
Gyromitrin



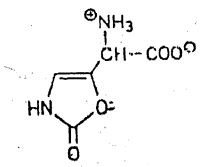
Muscarine



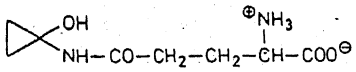
Ibotenic acid



Muscimol



Muscazone



Coprine

Chemical structure of mushroom toxins and hallucinogens

Other Amanitas that produce somewhat similar effects on ingestion are: *A. regalis*, considered by some to be only a brown variety of the fly-agaric, and *A. pantherina*, and *A. cothurnata*, all of which are known to contain ibotenic acid and muscimol.

The Psilocybes

The use of the fly-agaric by Siberian tribes finds a parallel in the ritual use of small mushrooms of the genus *Psilocybe* in Mexico. In pre-Spanish Mexico, they have long been in use to induce visions and have been called 'teonanacatl', literally, 'flesh of the Gods'. The human response to ritualistic administration of these hallucinogenic mushrooms by curanderos in Mexico is well described by Gordon Wasson from his own experience. To Aldous Huxley, the experience is: 'a state of affairs in which verbalizing and conceptualizing are in some sort by-passed. One can talk about the experience – but always with the knowledge that "the rest is silence"'. The latent period is short (15 to 120 min). Alteration of perception of time and space, a sense of elation and joy – or bliss, or alternatively (for others) anxiety and depression have been reported. The finale is a deep unconsciousness. These effects are comparable to those due to LSD (lysergic acid diethylamide) and have been amply described. Little had been known about the Mexican mushrooms until the ethnobotanical expeditions to Mexico by R. E. Schultes of Harvard University in the forties, and then followed, with the mushrooms which Schultes brought with him as the base, the pioneering and exciting studies of Gordon Wasson, mycologist Roger Heim of the Museum Cryptogamie in Paris, and Hoffman, biochemist with Sandoz in Basel on the ritual use, taxonomy and chemistry of these mushrooms, leading in essence to a remarkable synthesis of data. Roger Heim's invaluable monograph of these fungi is not only encyclopaedic but authoritative. Guzman's monograph on *Psilocybe* updates information. About 140 species of *Psilocybe* are known of which about 80 are hallucinogenic.

The largest number of hallucinogenic species are from Mexico. *P. aztecorum* var. *aztecorum* is believed to be the species used by the Aztecs. The species most popular or common in current use include *P. aztecorum* var. *aztecorum*, *P. zapotecorum*, *P. caerulescens* var. *caerulescens*, *P. cubensis* and *P. mexicana*. *P. cubensis* is coprophilous and comes up on cow dung: it is believed to have been introduced when the Spaniards brought in the cow from Africa. Several of the species can be grown in pure culture and can be cultivated. The active principle in several species is *psilocybin*, first isolated and characterized from *P. mexicana*, and also synthesized, by Hofmann. Psilocybin has the structure *N*, *N*-dimethyl-*H*-phosphoryloxytryptamine. The resemblance in structure to lysergic acid serves to explain the similarity in human response to psilocybin and LSD, already mentioned.

In nature, the unstable but related psilocin often occurs in small amounts along with psilocybin. In some species such as *P. baeocystis* and *P. semilanceata*, two other substances, also hallucinogenic, are known to occur: *baeocystin* and *nor-baeocystin*. Primarily, Psilocybe hallucinogens act as serotonin-antagonists: therefore they are used in psychotherapy and neurophysiological research. Psilocybin is active in man in very low doses: 4 mg taken orally causes wild intoxication, but 6–20 mg can elicit marked psychotropic response. Overdose can be dangerous in causing psychoses, leading to suicide mania.

Mushrooms in medicine – bioactive products

Apart from antibiotics, there are a number of substances that have been isolated from mushrooms and other basidiomycotina – substances that are variously labelled 'host defence potentiators', (HDPS), 'protein-bound polysaccharides' or 'polysaccharide–protein complexes' (PSPCs). The function of these is believed to lie in the toning up of defence mechanisms in the human system against a number of ailments. A great many of these are from polypores (Aphyllophorales) or other Basidiomycotina. At least about sixty species are traditionally used in Chinese medicine. These include species of *Ganoderma*, besides *Wolfiporia cocos*, *Polyporus mylittae*, *Fomitopsis officinalis*, *Coriolus versicolor*, *Fomes fomentarius* and *Bjerkandera fumosa*. *Tremella fuciformis* (Basidiomycotina, Tremellales) and *Coriolus versicolor* (Aphyllophorales) are used in Chinese folk medicine for their nutritive and tonic properties in the treatment of debility and exhaustion.

Ganoderma lucidum, *ling-chi* in Chinese folk medicine, is highly valued in China and in Japan for its medicinal properties and its ingredients, considered both a tonic and a panacea for a variety of ailments neither specific nor specifiable. An almost magical effect on intelligence, enhancing memory, and hearing, vision and smelling, has been attributed to this fungus. In practice, only the fruit bodies of *G. lucidum* and *G. japonicum* are used, the spores are discarded. During the past two decades there have been several studies on the chemistry of bioactive substances in *ling-chi* and the human response to these substances, but more specially their potential use in the control of a number of specific diseases such as cancer, diabetes and hepatitis. The spectrum of bioactive substances in *G. lucidum* includes organic germanium, polysaccharides (B-D-glucans), triterpenoids and adenosine.

Host-mediated anti-tumour polysaccharides of *G. lucidum* have been patented as immunopotentiators in the treatment of cancer. Besides, *G. lucidum*, *G. tsugae* and *G. boninense* are included in Japanese and US patents. Also, over one hundred novel, highly oxygenated, lanosa-type triterpenoids have been reported from fruit-bodies and mycelium of *G. lucidum*: their chemical structure has been determined. These include: ganoderic acids, gan-

oderemic acids, ganodermic acids, ganoderal, lucidernic acids and lucidone and ganoderols. It is tempting to speculate if these diverse triterpenes could explain the multiple beneficial effects for which *ling-chi* is reputed.

A freeze-dried powder containing mycelium of *G. lucidum* has been shown to bring down blood sugar levels in experimental diabetic rats. Three hypoglycaemic principles, *ganoderans* A, B and C have been isolated from fruit-bodes: these are characterized as peptidoglycans. Of these, ganoderan B is considered to be the major one.

Extracts of the fungus have also been shown to be hepato-protective: apart from liver regeneration, beneficial effects against hepatic necrosis and hepatitis have been noted. *Ganodosterone* from *G. lucidum* is a liver-protectant and stimulates liver function. Similarly, ganoderic acids T, S and R, from the mycelium of the fungus are hepato-protective. Hepato-protective triterpenoids have also been isolated from *G. tsugae*: lucidone A, lucidenol, ganoderic acid B, and ganoderic acid C2.

Ganoderma is cosmopolitan; some species cause serious decay in wood; some are root-disease pathogens of a variety of plantation crops and are economically important. Some species have the ability to selectively degrade lignin and so are a potential source of lignin-degrading enzymes useful in the bioconversion of lignin in agricultural and other wastes. The new wave of interest in *Ganoderma* has prompted studies on taxonomy, distribution, and even speciation. The low nucleotide variation observed in the 25 S rDNA and the very high transition bias in the region suggest that the genus is young and is still evolving.

Lentinan from the edible mushroom, *Lentinus edodes* has been shown to have distinct anti-tumour properties and has also potential use in enhancing host resistance to a variety of infections, including HIV-1 or AIDS. *Lentinus edodes* is remarkably effective in lowering levels of lipids in the blood: the hypolipidemic effects are due to a secondary metabolite identified as 2(R), 3(R)-dihydroxy-4-(I-adenyl)-butyric acid which has the trivial name *eritadenine* (=*lentysine, lentinacin*). A lowering of serum cholesterol levels in man has also been reported.

Mushrooms: antiquity, speciation

The oldest known fossil mushroom species is perhaps *Coprinites dominicana* described as trapped in amber about 40 Mya. Mushrooms are believed to have radiated in the Cretaceous about 130 Ma, after the great rise of the angiosperm land flora. Ectomycorrhizal fossil mushrooms date back to the same time.

In the case of *Pleurotus* (the oyster mushroom) compatibility tests have shown the existence of several distinct intersterility groups (biological species). Recent work on speciation, from phylogenetic analysis of ribosomal DNA sequences among these intersterility groups, indicates that several species lineages evolved quite early: species lineages of more recent origin appear to be restricted in distribution, viz. to the northern hemisphere. Older lineages, on the other hand, have a distribution world-wide. The explanation for this pattern of distribution may well be that the old lineages were extant before the break-up of 'continents' from drift, and new lineages arose in the northern hemisphere following continental drift. On the other hand, one might also postulate dispersal as the primary factor governing the current distribution pattern. Whichever of these two postulates we accept, it is clear that older lineages have a wider distribution and more recent lineages have relatively narrower distribution, some thing akin to Willis's Age and Area hypothesis. Genetic divergence appears to be linked to allopatric speciation within populations from different continents in the northern hemisphere.

This contribution is but a feeble attempt to capture and present for the reader something of the beauty, diversity, and relevance of mushrooms. The literature on the subject is enormous and an account like this cannot be anything but incomplete. The list of selected references appended here should, however, help the reader to delve deeper into the subject. And, it would give me great joy if what I have written and presented provokes curiosity, further probing into 'mushroom science' and, most important, evokes excitement, fascination and wonder.

1. Alder, A. E., *Z. Pilzk.*, 1960, **38**, 65–73.
2. Bakshi, B. K., *Indian Polyporaceae* (on Trees and Timber), Indian Council of Agricultural Research, 1971, New Delhi
3. Berbee, M. L. and Taylor, J. W., *Can. J. Bot.*, 1993, **71**, 1114–1127.
4. Bresinsky, A. and Besl, H., *A Colour Atlas of Poisonous Fungi* (Transl. by N.G. Bisset), Wolfe Publishing Ltd., 1990.
5. Buchanan, P. K., Hseu, R. S. and Moncalvo, J. M. (eds), *Ganoderma – Systematics, Phytopathology and Pharmacology*, 1995 (in press).
6. Buller, A. H. R., *Researches on Fungi, I–VI*, Longmans Green, London, 1909–1934.
7. Chang, S. T., Buswell, J. A. and Chiu, S. W. (eds), *Mushroom Biology and Mushroom Products*, The Chinese University Press, Hongkong, 1994.
8. Clayton, N., in *Secondary Metabolic Products of Selected Agarics* (eds. Moore D., *et al.* cited below.), 1985, pp. 561–579.
9. Daniels, E. G., Kelly, R. B. and Hinman, J. W., *J. Am. Chem. Soc.*, 1961, **83**, 3333–3334.
10. Faulstich, H., *Frontiers in Mycology* (ed. Hawksworth, D. L.), CAB International, 1991, pp. 211–223.
11. Frankland, J. C., Hedger, J. N. and Swift, M. J. (eds) *Decomposer Basidiomycetes: Their Biology and Ecology*, Cambridge University Press, Cambridge, 1982.
12. Grover, S. (ed.), *Letters of Aldous Huxley*. Harper & Row, 1969, New York.
13. Guzman, G., *The Genus Psilocybe*, J. Cramer, Vaduz, 1983.
14. Harley, J. L., *Proc. Indian Acad. Sci. (Plant Sci.)*, 1985, **94**, 99–109.
15. Harley, J. L. and Smith, S. E., *Mycorrhizal Symbiosis*, Academic Press, London, 1983.
16. Hawksworth, D. L., Sutton, B. C. and Ainsworth, G. C., *Ainsworth & Biosby's Dictionary of Fungi*, CMI, Kew, 1983.
17. Heim, R., *C.R. Acad Sci.*, 1956, **242**, 965–968, 1389–1395.
18. Heim, R., *C.R. Acad. Sci.*, 1957, **244**, 659–700.
19. Heim, R., *Rev. Mycol.*, 1957, **22**, 58–79.
20. Heim, R., *Termites et Champignons. Les Champignons Termitophiles d'Afrique Noire et d'Asie Meridionale*. Editions Boubee, Paris, 1977.

21. Heim, R., *Les champignons toxiques et hallucinogenes*, 2nd edn. Boubee, Paris, 1978.
22. Jennings, D. H. and Rayner, A. D. M. (eds), *The Ecology and Physiology of the Fungal Mycelium*, Cambridge University Press, Cambridge, 1984.
23. Jong, S. C. and Birmingham, J. M., *Adv. Appl. Microbiol.*, 1992, **37**, 101–134.
24. Kaneda, T. and Tokuda, S., *J. Nutrition*, 1966, **90**, 371–376.
25. Last, F. T. and Fleming, L. V., *Proc. Indian Acad. Sci. (Plant Sci.)*, 1985, **94**, 111–127.
26. Lincoff, G. and Mitchel, D. H., *Toxic and Hallucinogenic Mushroom Poisoning*, Van Norstrand Reinhold & Co., NY, 1977.
27. Liu, J.-W., Beelman, R. B., Lineback, D. R. and Speroni, J. J., *J. Food Sci.*, 1982, **47**, 1542–1544.
28. Manjula, B., *Proc. Indian Acad. Sci. (Plant Sci.)*, 1983, **92**, 81–213.
29. Moncalvo, J. M., Wang, H. H. and Hseu, R. S., *Mycologia*, 1995, **87**, 223–238.
30. Moore, D., Casselton, L. A., Wood, D. A. and Frankland, J. C. (eds), *Developmental Biology of Higher Fungi*, Cambridge University Press, Cambridge, 1985.
31. Moser, M., *Proc. Indian Acad. Sci. (Plant Sci.)*, 1985, **94**, 381–386.
32. Natarajan, K. and Raman, N., *South Indian Agaricales*. Cramer, Vaduz, 1983.
33. Poinar, G. O. and Singer, R., *Science*, Washington D.C., 1990, **24**, 1099–1101.
34. Ramsbottom, J., *Mushrooms and Toadstools*, Collins, London, 1953.
35. Ross, A. E., Nagel, D. L. and Toth, B., *Food Chem. Toxicol.*, 1982, **20**, 903–907.
36. Rumack, B. H. and Salzman, E. (eds), *Mushroom Poisoning: Diagnosis and Treatment*, CRC Press, Florida, 1978.
37. Schultes, R. E. and Hofmann, A., *Plants of the Gods: Origins of Hallucinogenic Use*, Hutchinson, London, 1979.
38. Singer, R., *Mushrooms and Truffles. Botany, Cultivation and Utilization*, Leonard Hill, London, 1961.
39. Singer, R., *The Agaricales in Modern Taxonomy*, Cramer, Vaduz, 1975, 3rd edn.
40. Singer, R., *Marasmieae (Basidiomycetes, Tricholomataceae)* Flora Neotropica. Monograph No. 17. New York Botanical Garden, 1976, New York.
41. Speroni, J. J., Beelman, R. B. and Schisler, L. C., *J. Food. Prot.*, 1983, **46**, 506–509.
42. Subramanian, C. V., *Curr. Sci.*, 1991, **60**, 290.
43. Suzuki, S. and Oshima, S., *Mushroom Sci.*, 1974, **9**, 463–467.
44. Toth, B., *J. Toxicol. Environ. Health*, 1979, **5**, 193–202.
45. Tullos, R. E., Ovrebo, C. L. and Halling, R. E., *Mem. N.Y. Bot Gdn.*, 1992, **66**, p. 46.
46. Vilgalys, R. and Sun, B. L., *Proc. Natl. Acad. Sci. USA*, 1994, **91**, 4599–4603; 7832.
47. Wasson, R. G., *Soma, Divine Mushroom of Immortality*. Harcourt Brace & World, N.Y., 1968.
48. Wieland, Th. *Peptides of Poisonous Amanita Mushrooms*, Springer, Berlin, 1986.
49. Zhao, Ji-Ding, *The Ganodermataceae in China*. J. Cramer, 1989, Berlin.
50. Zhao, Ji-Ding. and Zhang, Y-Q., *The Polyporaceae of China*. J. Cramer, 1992, Berlin.

Note:
Authorities for Latin binomials are not given for ease of reading. These will be found in the monographs cited. Photo credits: Figures 1, 2, 5 courtesy K. Natarajan; 7, 10, 11 courtesy Chen, Hsinchu.

RESEARCH ARTICLE

Role of fluorescence microscopy in the assessment of Indian Gondwana coals

Bhagwan D. Singh

Birbal Sahni Institute of Palaeobotany, GPO Box 106, Lucknow 226 001, India

When a light of short wavelength excites organic matter, light of relatively longer wavelength is emitted from it and this phenomenon is known as autofluorescence. The coal maceral analysis under fluorescence mode (blue light/UV light excitation), therefore, has been found to be best suited to properly identify, characterize and quantify hydrogen-rich macerals. Utilizing this technique, macerals like bituminite, fluorinite and exsudatinite were recognized for the first time. Certain other macerals – alginite and liptodetrinite, normally mistaken for mineral matter under routine petrographic analysis, were also identified. Fluorescence microscopy, thus, not only added to the overall tally of liptinite group of macerals in Indian Gondwana coals, but also to their quantity. In addition to this, recognition of fluorescing vitrinite (perhydrous vitrinite) significantly contributed to the abundance of hydrogen-rich microconstituents for these coals.

FLUORESCENCE microscopic study of solid fossil fuels (coals/lignites) and kerogen (in sedimentary rocks) has been found quite successful particularly in identification, characterization and quantitative assessment of hydrogen-rich macerals, especially the liptinite (or exinite) group of macerals. It is a useful technique for determination of rank or maturity and even better suited in cases where vitrinite/huminite macerals are not in sufficient quantities or not suited for reflectivity measure-

Presented at 'Contemporary Research Trends in Palaeobotany' to commemorate The Palaeobotanical Society International Medal Award function, BSIP, Lucknow, January 12, 1993.

ments, e.g. sapropelic coals, bog-head coals, oil shales, kerogen, etc. It affords a quantitative estimation of spectral fluorescence emission which is characteristic of both the individual maceral as well as rank/maturity and enables easy and reliable distinction of liptinites.

The liptinite maceral group, which is chemically distinguished from its two precedents – inertinites (relatively carbon-rich) and vitrinite (relatively oxygen-rich) by its relatively higher hydrogen content, is particularly heterogeneous in its biological origin. Because it mixes organs (or fragments of organs) and plant organisms: these are, on the one hand, spores–pollen grains (sporinite), cuticles (cutinite) and certain epidermal tissues, plant secretions – resins (resinite), essential oils, etc. and, on the other hand, the immense class of algae (alginite). Certain other liptinite macerals originate during coalification process, e.g. liptodetrinite, fluorinite, bituminite; derived from bacterial substances and decomposition products (Table 1), besides a secondary maceral – exsudatinite. Plants lipids, proteins, cellulose and other carbohydrates are characteristic source materials of liptinites[1].

Evidently, various liptinite macerals having different modes of occurrences and chemical properties behave differently under fluorescence mode and their fluorescence characteristics are correlatable with rank stages. They are all characterized by a fluorescence which takes on various colouration according to their nature and their rank; the intensity of this fluorescence decreases with the increasing rank and disappears towards 1.0 to 1.3% R_0 max[3,4]. Since the fluorescence property of organic matter or maceral is the reflection of its overall chemical characteristics, the technique helps in understanding the changing pattern of macerals through various coalification stages.

Gondwana coals are normally rich in mineral matter/ash contents. Also their liptinite contents recorded under normal reflected light is quite low (usually 1–15%). In fact, high inherent clastic minerals intimately associated with these macerals in the coals tend to mask them, at times completely, rendering their identification difficult or even impossible under normal incident light. Earlier, only macerals like sporinite, cutinite, suberinite and resinite were subnormally recorded, whereas liptodetrinite, alginite, fluorinite, bituminite, etc. were normally mistaken for mineral matter. It is primarily because of this reason that mineral matter-rich Lower Gondwana (Permian) coals of India were and are still considered to be poor in liptinite contents.

In the past few decades it has been realized that chemical properties of Permian coals, viz. carbonization, swelling index, liquefaction, etc. do not correspond well with the biopetrological results under normal reflected light. The explanation was not possible till the use of fluorescence microscopy came into routine service, by which hydrogen-rich macerals quantities show a manifold increase in the same coals. It is the intention of this communication to highlight the advantage of fluorescence microscopic method in evaluation of certain Lower Gondwana coal seams from Raniganj, Singrauli and Rajmahal fields of India.

Table 1. Classification of coal maceral groups/macerals and their source material. (ref. 2, modified after ref. 1)

Group	Maceral	Origin
Vitrinite	Telinite	Cell walls (may be resin impregnated or lignified) rich in cellulose.
	Telocollinite	Gelified plant tissues.
	Desmocollinite	Gelified humic detritus (amorphous).
	Vitrodetrinite	Humic detritus.
	Corpocollinite	Cell filling (oxidation/condensation products of tannins or biochemically gelified humic matter).
Liptinite or Exinite	Sporinite	Spores and pollen exines of higher plants.
	Cutinite	Outer layers of leaves, needles, shoots, thin stems of higher plants.
	Suberinite	Suberinized cell walls (bark or cork cells).
	Resinite	Resin, balsam, copal, latex, wax, oils and fats from higher land plants.
	Fluorinite	Essential oils of higher land plants.
	Alginite	Colonial and unicellular algae.
	Bituminite	Anaerobic biodegraded products (amorphous microbial, algal and/or faunal remains).
	Chlorophyllinite	Chlorophyll.
	Liptodetrinite	Biodegraded or detrital liptinite macerals and/or phytoplanktons and algae.
	Exsudatinite	Secondary exudates.
Inertinite*	Fusinite	Cell walls (charred, oxidized or fungus-infested).
	Semifusinite	Cell walls (partly charred, oxidized, or fungus-infested).
	Sclerotinite	Fungal hyphae, mycelia, spores and sclerotia.
	Macrinite	Amorphous gel (oxidized/microbial metabolic products).
	Inertodetrinite	Detritus of above inertinite macerals.
	Micrinite	Secondary relics of oil generation (mostly).

*A small part of inertinite originates from melanin-rich plant and animal material – 'primary inertinite'. A greater part attains its inertinitic properties during early coalification process – 'rank inertinite'.

Materials and method

In principle, irradiation or excitation with a light of short wavelength, routinely blue or ultra-violet light, on organic matter or maceral causes emission of fluorescence colours and this phenomenon continues till stoppage of excitation – 'auto-fluorescence'. During auto-fluorescence, emitted colour and intensity changes with time and this dual change is known as 'alteration effect'. Alteration of fluorescence colour and intensity are typical of a particular organic matter or maceral at a given

rank or maturity stage. Based on these criteria, various types of organic matter/macerals are identified and characterized.

Maceral content was determined under blue light excitation (filter block H3:420-490 nm, violet-blue excitation) on Leitz MPV-3 unit using 150 W ultra high pressure xenon lamp (to initiate fluorescence) as source of illumination, 25× NPL Fluotar objective (oil) with a 0.75 numerical aperture and fluorescence-free immersion oil. The reported results, based on 500–1000 point counts per sample, are expressed as volume per cent of the various fluorescing and non-fluorescing macerals on a mineral matter-free basis.

A total of 192 particulate coal pellets prepared by cold embedding in epoxy resin were analysed by coal petrographic procedures. Some pertinent details about number and nature of samples, name of coal seams and their respective coalfields are given in Table 2. The coals range in rank between sub-bituminous A to high volatile bituminous C–A stages (R_0 max. 0.40–0.88% in oil).

Table 2. Details of coal samples studied from some Lower Gondwana coalfields of India

Late Permian – Raniganj Formation
Raniganj Coalfield (Damodar Basin), West Bengal
Three bore-hole sections from:
Girmint block (bore-hole TAD GRT/0160/CM-009)
Seam nos. (bottom to top): I, II, III bot., III top, IV, Bankshimulia, V, VI, Century, VII, VIII bot., VIII top, Kusudanga, IX, XA and X (16 samples)
Dhemo Main block (bore-hole DMM/0030/CM-006)
Seam nos. (bottom to top): VA, V, VI bot., VI mid. VI top, VIIA, VII, VIII bot, VIII top, IX and X (11 samples)
Shyamsunderpur block (bore-hole SS-5)
Seam nos. (Bottom to top): IV, V, VI, VIIA, VII, VIII bot., VIII mid., VIII top and IX (9 samples)
Singrauli Coalfield (Son Basin), Madhya Pradesh
One bore-hole section of Jhingurdah seam from:
Jhingurdah colliery (bore-hole NCSJ-4)
Top seam (41 samples)
Bottom seam (6 samples) (47 samples)
Early Permian – Barakar Formation
Singrauli Coalfield (Son Basin), Madhya Pradesh
Five sections of Turra seam from:
Kakri (10 samples), Bina (11 samples), Jayant (8 samples),
Gorbi (16 samples) collieries and bore-holes NCSM-3 (Moher block-15 samples) and CMSA-111 (block-B-7 samples) (67 samples)
Chuperbhita Coalfield (Rajmahal Basin), Bihar
One bore-hole section from:
Chuperbhita basin (bore-hole RCH-3)
Seam nos. (bottom to top): I (samples), II (9 samples), III (6 samples) IV (4 samples), V (3 samples), VI (3 samples), VII (3 samples) and VIII (9 samples) (42 samples)

Caution

It must be mentioned here that in identifying organic matter/coal macerals in fluorescence mode, one should have proper morphological knowledge. There are some minerals, viz. calcite, siderite, fluorite, a variety of quartz, etc. which fluoresce and can be mistaken for resinite and cutinite macerals. Certain translucent minerals, such as clay, because of internal reflection/refraction appear green, yellowish-green to yellow colours giving a false impression of being resinite. Beside this, embedding medium being a chemical substance, does react with embedded coal particles producing yellowish-green – yellow to yellowish-orange colours and, thus may add to the tally of resinite.

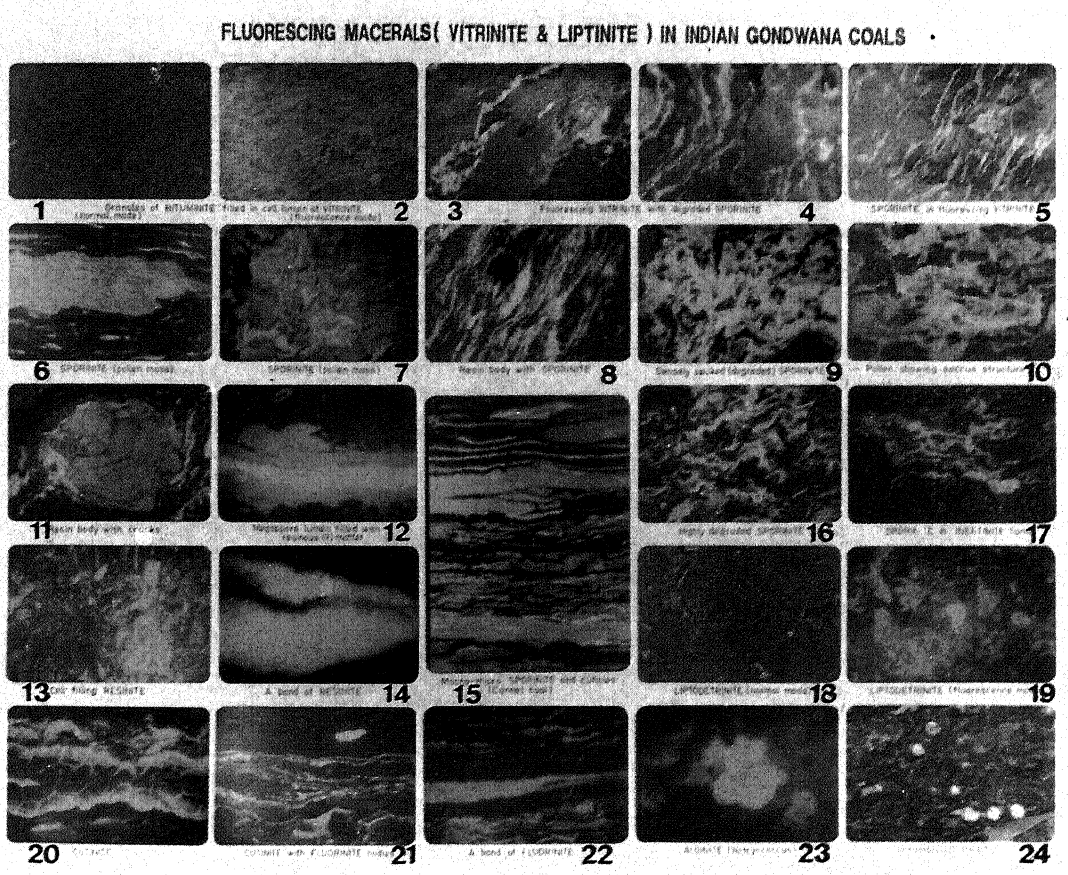
The quantitative evaluation of macerals under fluorescence mode may tend to give an over estimate of fluorescing matter. This is because of the fact that some of the short wavelength light penetrates into underlying coal and, thus generates three-dimensional image of macerals below the polish surface of the grain mount. Therefore, quantitative estimation of macerals, especially the liptinite macerals, in mineral matter-rich coal samples may represent over estimated results. However, in clean coals this factor may not apply at all.

Characteristic features of fluorescing macerals

Perhydrous vitrinite and all the macerals of liptinite group were the fluorescing macerals in the Lower Gondwana coals of the Raniganj, Singrauli and Rajmahal coalfields. Corpocollinite and a fraction of telocollinite macerals of vitrinite group and macerals of inertinite group were non-fluorescing in these coals.

Sporinite is derived from the outer cell walls of spores and pollen. It fluoresced with orangish-yellow, yellowish-orange, brownish-yellow to yellowish-brown colours in studied Gondwana coals. The sexinal part of the spores and pollen fluoresced with relatively brighter colour than the nexinal part bordering streaky slit- or thread-like lumen. Both tenui- and crassi-sporinites were present along with sporadic to common macrosporinite (Figures 12, 15). The microsporinite in coals was either sparsely distributed (Figures 3, 4) or was packed densely in thin to thick microbands (Figures 5, 8, 15, 22), especially in dull coals. Densely packed sporinite as thick microbands were frequent in Singrauli[2] and Rajmahal coals, where a part of dull bands evidently developed into semi-cannel or cannel coals. This is the first record of the presence of semi-cannel/cannel coal bands in India. The sporinite-rich bands were usually poor in the inertinite macerals which occurred randomly as pieces and fragments (Figure 17).

The Lower Gondwana coals commonly contained gymnospermous pollen-mass or pollen in sporangium with or without sporangial covering/wall in unreleased stage (Figures 6, 7). In normal light, such dark grey



Figures 1, 2. Granules of bituminite filled in cell lumens of vitrinite, Singrauli Coalfield (Figure 1 under normal mode); **Figures 3, 4.** Fluorescing (perhydrous) vitrinite with degraded sporinite in Raniganj (3) and Singrauli (4) coals; **Figure 5.** Sporinite in fluorescing vitrinite, Rajmahal Coalfield; **Figures 6, 7.** Sporinite (pollen mass) in Singrauli coals; **Figure 8.** Resin body with sporinite in Singrauli coals; **Figure 9.** Densely packed (degraded) sporinite in Singrauli coals; **Figure 10.** Pollen showing saccus structure (sporinite) in Singrauli coals; **Figure 11.** Resin body with cracks in Singrauli coals; **Figure 12.** Megaspore lumen filled with resinous (?) matter in Rajmahal coals; **Figure 13.** Cell filling resinite in Singrauli coals; **Figure 14.** A band of resinite in Rajmahal coals; **Figure 15.** Cannel coal, macro-, microsporinite and cutinite in Singrauli coals; **Figure 16.** Highly degraded sporinite in Singrauli coals; **Figure 17.** Sporinite in inertinite band of Rajmahal coals; **Figures 18, 19.** Liptodetrinite in Rajmahal coals. (Figure 18 under normal mode); **Figure 20.** Cutinite showing cuticular ledges, Singrauli coals; **Figure 21.** Cutinite with fluorinite bodies in Singrauli coals; **Figure 22.** A band of fluorinite in Rajmahal coals; **Figure 23.** Alginite (*Botryococcus*) in Rajmahal coals; **Figure 24.** Framboidal pyrite in Singrauli coals under normal mode. [All photomicrographs were taken on polished surface under incident blue light excitation (except nos. 1, 18 and 24) using oil immersion]

densely clustered sporinite was earlier misidentified either as calcite mineral or resinite maceral by others. The individual pollen of the pollen-mass are tightly packed together so that their outline is generally obscured but their lumen can invariably be recognized.

During the investigation, it was observed that sporinite showed wide variety of preservational stages from well-preserved to highly degraded and fragmented (Figures 9, 10, 16). During biodegradation the pollen appears to have disintegrated into smaller pieces by granulation and ultimately got reduced to fine granular mass of liptodetrinite. Mechanical breaking of pollen with or without biodegradation also led to liptodetrinite formation. Another form of sporinite degradation noticed was by *in situ* pyrite growth where pyrite crystals tended to grow inside the pollen saccus, rupture it by blistering the saccus wall, thus facilitated further biodegradation and mechanical breaking.

Cutinite is derived from cuticles and cuticular layers which occur at the surface of leaves, twigs, and other areal parts of plants as a protection from dessication. It showed brownish-orange to orangish-brown colours with relatively weaker fluorescence intensity than the sporinite in these coals. Both tenui- and crassi-cutinites were observed generally in well-preserved state (Figures 15, 20). Cutinite enclosing certain tissue structure, such as leaf sections, was normally associated with fluorinite which appear as irregular amorphous bodies of various dimensions (Figure 21).

Fluorinite originates from essential oils and lipid-rich cell inclusions of various sources. It occurs as blobs or blotches, streak and lensoid bodies or as thin bands and

Table 3. Maceral composition (mean and range) and reflectance value of the coal seams from the studied fields of India

	Raniganj Coalfield	Singrauli Coalfield		Rajmahal Coalfield
	Seams I–X	Jhingurdah seam	Turra seam	Seams I–VIII (Chuperbhita)
Under normal incident light (volume %)				
Vitrinite	59.90	54.23	37.58	36.57
	41.4–77.0	25.0–74.0	3.6–76.0	5.0–80.8
Liptinite	5.24	3.32	5.66	4.83
	1.0–11.0	1.0–5.4	1.0–13.6	1.0–12.2
Inertinite	14.51	21.98	25.20	36.22
	4.2–23.2	11.2–26.6	6.0–60.6	8.0–72.0
Mineral matter	20.35	20.47	21.56	22.38
	8.0–40.6	10.4–25.2	4.0–36.4	3.6–49.4
R_0 max. (% in oil)	0.56–0.88	0.54–0.60	0.43–0.60	0.43–0.57
Under blue light excitation (mineral matter-free %)				
Vitrinite	49.08	44.16	29.03	30.33
(fluorescing)	20.8–64.0	19.0–57.0	6.0–54.0	2.0–80.2
Liptinite	20.98	25.10	32.13	33.80
	14.3–27.4	13.5–53.1	14.5–57.4	5.4–66.2
Sporinite	10.29	15.12	21.49	24.28
	4.9–15.5	7.2–39.7	9.6–38.3	2.0–58.0
Resinite	1.60	0.83	1.08	0.32
	0.4–3.5	0.0–2.9	0.0–2.1	0.0–3.2
Other liptinites	1.24	1.07	3.17	0.18
	0.0–3.8	0.2–3.1	0.5–6.5	0.0–2.6
Liptodetrinite	7.85	8.08	6.39	9.02
	4.5–13.2	3.8–15.7	1.2–12.0	1.6–20.0
Vitrinite + inertinite	29.94	30.74	38.84	35.87
(non-fluorescing)	19.2–64.9	23.8–46.3	24.0–59.2	12.0–72.0

having strong fluorescence intensity (of bright orangish-yellow to yellowish-orange fluorescence colours; Figures 21, 22). Such fluorinite bodies can be easily mistaken for resinite even in fluorescence mode, unless their rapid alteration is ascertained. Because of indefinite morphology, amorphous nature and characteristic colour (dark grey), the maceral fluorinite is unidentifiable under normal light and is mistaken for voids or translucent minerals.

Suberinite is derived from suberin layers of corkified cell walls, mainly in tree barks. Suberin is formed also at the surface of roots and fruits, acting as a protection against dessication. The maceral suberinite fluoresced with weak orangish-brown colour and was only sporadically observed in these Gondwana coals.

Alginite represented by *Botryococcus* was commonly observed occurring discretely or in clusters in well-preserved to highly degraded conditions. Well-preserved alginite was not observed in the Late Permian coal seams. It fluoresced more strongly (with orange colour) than other liptinites and showed rapid alteration. In degraded condition it was recognized by its characteristic fluorescence colour, intensity and relict morphology. Alginite appearing as dark grey and black coloured discrete globular or small lumpy bodies with granular texture under normal reflected light has been misidentified as mineral lumps. Presence of *Botryococcus* in Singrauli (persistent[2]) and Rajmahal (common) fields is the first authentic record from the Palaeozoic coals (Figure 23). It seems that these coals contain a fair amount of alginite, most of which was later degraded and incorporated in perhydrous vitrinite. Contribution of alginite to humic coals in fair amounts has already been recognized by Stach *et al.*[5]

Resinite is derived not only from resins but also from balsams, latexes, fats and waxes. It is characterized by variable fluorescence intensity and colour than any other liptinite macerals in a given coal. It occurred as discrete oval, elliptical or elongated bodies with regular or irregular margins of various dimensions (usually small) in Gondwana coals (Figures 8, 11, 13, 14). Macerals such as collinite in telocollinite/desmocollinite as thin micro strips, streaks or stringers in bi- and trimaceral associations have almost identical light to dark grey colours as that of resinite and cutinite under normal light. It is for this reason that there is a possibility of misidentification of certain vitrinite maceral as resinite. In addition to this, transluscent white or pale yellow clay blobs or blotches normally non-aligned with the surrounding macerals, as has already been mentioned, are likely to be mistaken as green or yellow resinites in fluorescence mode.

In fact, green or yellowish-green resinites were not recorded in these coals. Greenish-yellow resinite was rare and that fluoresced with yellow colour was sporadic as expected according to rank of these Gondwana coals.

Greenish-yellow to orange fluorescing resinites were generally translucent, whereas brown resinite appeared opaque. With increase in rank, it has been proved in Indian coals and lignites[6] that resinite fluorescing with green, yellowish-green or greenish-yellow and yellow colours gradually become less common, being overtaken by orange and brown colours (including transitionary colours).

Liptodetrinite is a common product of anaerobic microbial degradation and disintegration of liptinite macerals, chiefly sporinite, possibly alginite and certain amount of cutinite in these coals. Participation of resinite maceral in liptodetrinite formation was not recognized in studied coals. Because of its highly degraded nature and lack of morphology, the maceral is normally mistaken for mineral matter-rich patch under normal incident light (Figures 18, 19). Its recognition is routinely possible under fluorescence mode only.

Bituminite is the product of anaerobic biodegradation formed in coals as a result of 'bituminization' during biochemical gelification stage between vitrinite R_0 max. >0.40 to 0.50% (ref. 7). It is an amorphous (unstructured) maceral of liptinite group with brownish fluorescence colour of weak intensity. This maceral is properly recognized only under fluorescence mode. Normally, the bituminite formed is pervasively impregnated in the pore spaces of vitrinite rendering it perhydrous in nature (Figures 1, 2). Therefore, most of the bituminite present in a coal is accounted for alongwith perhydrous vitrinite (imparting weak to moderately weak reddish-brown fluorescence colour) including desmocollinite groundmass. Sometimes granules of bituminite were observed packed densely inside the original cell lumens of vitrinite separated by non-fluorescing cell walls. Presence of alginite in studied coals indicate that significant amount of bituminite was generated by alginite as well.

Exsudatinite originates as fluid substances expelled from lipid-rich parent material (in vitrinite, fluorinite, resinite, etc.) during 1st coalification jump. Appearing as migrated bitumen, in coals it fills open cleats, fissures and pores. The formation of exsudatinite is often associated with oil expulsions in the form of smear films and as droplets exuding/oozing from cleats in vitrinite. These oily exudates are not considered as maceral by Teichmüller[8]. In studied coals, exsudatinite and surface films were only sporadic in occurrence and fluoresced with weaker intensity than sporinite and some of the cutinites in orangish-brown or brownish-orange colours. In low rank coals, exsudatinite and oily droplets appear brown or dark grey in colour under normal light, whereas the smear films often show rainbow colour bands. This secondary maceral is properly recognized only under fluorescence mode.

Biopetrographic composition of coals

The coal seams of Barakar Formation from Rajmahal and Singrauli fields are normally rich in inertinite (6.0–72.0%) and the vitrinite ranges between 3.6 to 80.8%. The Raniganj coal seams (from Singrauli and Raniganj fields), on the contrary, contain high vitrinite (25.0–77.0%) and relatively low inertinite (4.2–26.6%). These Gondwana coals are characterized by persistently low (1.0–13.6%) and randomly distributed liptinite contents (Table 3). Associated minerals are usually fine black granular argillaceous matter, besides siderite, calcite and pyrite minerals, commonly found dispersed and intimately associated with organic constituents.

Under blue light excitation, the coals of both Barakar and Raniganj Formations were found to contain appreciably high amounts of liptinite macerals (7.0–66.2%) and perhydrous vitrinite (6.0–80.2%). However, the coal seams of the Barakar Formation are characterized by higher proportions of liptinite macerals, especially the sporinite (4.9–58.0%) than the coal seams of the Raniganj Formation (4.8–39.7%), whereas the coal seams of the latter formation have higher proportion of perhydrous vitrinite (19.0–64.0%) than those of the former (2.0–80.2%). The liptinite macerals in these coals are formed chiefly by sporinite (4.8–58.0%) and liptodetrinite (1.2–20.0%). Cutinite, suberinite, resinite, exsudatinite and fluorinite are other liptinite macerals together occurring in only subordinate amounts. Maceral bituminite being associated with perhydrous vitrinite and liptodetrinite was not recorded separately. Presence of maceral alginite was commonly recorded (up to 0.8%) in the Barakar coal seams.

It must be mentioned here that Chakrabarti[9], Chaudhuri and Ghose[10], and Saxena *et al.*[11] carried out fluorescence microscopy of certain Indian Gondwana coals. However, their petrographic data in certain specific aspects do not correspond well with those obtained presently. Therefore, it has been deemed necessary to explain the reasons of the discrepancies between the two data sets.

Chakrabarti[9] has evaluated representative coal samples (seams) from various Gondwana fields and remarked that 'higher-volatile bituminous coals (VM >40%) with reflectance values varying from 0.45–0.65% do not show any fluorescence property in primary structured constituents like vitrinite and fusinite'. Absence of fluorescing (perhydrous) vitrinite in this rank range is surprising, even coals having abundant liptinites. However, desmocollinite and a fraction of telocollinite (of vitrinite group) were presently found to be fluorescing with dull reddish-brown to dark brown colours, although with weak intensity.

The bituminite recorded by Chaudhuri and Ghose[10] (Figure 7) does not conform to its characteristic morphology. It is more like a resinite, whereas 'greenish-yellow' colour of the associated exsudatinite (in same figure) is rather surprising for high volatile bituminous coal rank. They also recorded crack-filling fluorinite. Actually it is not a crack-filling maceral. As is clear from their figure (8a, p. 248) it is an embedding medium between two coal particles and a crack in one of them.

Saxena *et al.*[11] carried out analysis of Raniganj coal seams (R-I – R-X) from the Raniganj field. They have not, however, provided information about the location and sequence of samples, lateral representation of coal seams and fate of split seams, besides coal quality and rank. Therefore, generalization made on the coal seams (for the coalfield) by them is unacceptable. Absence of cutinite under blue light excitation (Table 2, p. 61), but their presence in certain seams under normal light (Table 4, pp. 66, 67) is surprising. More than 5–10% (m.m.f) of resinite and 65 to above 69% (m.m.f.) of perhydrous vitrinite contents reported by Saxena *et al.*[11] (Table 2) is also abnormally high. Such high quantities of perhydrous vitrinite is not recorded even in Tertiary coals of Assam which are best suited for liquefaction[12]. High resinite content (5–10%) in coal seams R-II, R-III, R-VII to X recorded by Saxena *et al.*[11] is because of wrong identification of macerals collinite, desmocollinite, corpocollinite and sporinite and calcite, siderite and a certain fraction of clay minerals as resinite. The following are some examples from the illustrations of Saxena *et al.*[11] to support my contention:

(1) Plate 1, Figures 1, 2, 4, 5: the resinite (R) shown are collinite-filled lumens of telocollinite.

(2) Plate 3, Figure 2: resin (R) fillings and bitumen (B) shown are sporinite and calcite respectively. Figure 4: is not resin fillings in irregular spaces but sporinite in desmocollinite.

(3) Plate 4, Figures 1, 2, 6: are not resinite (R) but anisotropic siderite whose dark grey portions were identified as resinite. Refer siderite concretion in Plate 5, Figure 1. Figures 3, 4: are not resinite but corpocollinite and desmocollinite (in trimacerite) respectively. Figure 5: is calcite not resin (R) content filled in elliptical shape.

(4) Plate 5, Figure 3: is a pollen-mass (sporinite) not resin (R) fillings along the cracks. Note sporinite lumens are clearly discernible.

(5) Plate 6, Figures 2,3: illustrate green resinite (GR) which, in fact, are voids of fusinite lumen filled with embedding medium. Green to orange resinites do not fluoresce with such weak intensity. Figure 4: is not yellow resinite (YR) but oily droplets (exsudatinite) oozing out from a crack in vitrinite. Figure 5: illustrates two identical patches one (arrow pointing orange colour thread-like structures which is, in fact, lumen) recognized as sporinite (S) and the other (arrow pointing coalesced sporinite walls) as yellow reactive resin (YR)/hydrocarbon. In fact, both of them are pollen-mass (sporinite). Sporinite wall can never become resinite. Figure 6: is not green reactive resinite (GR) filled in fusinite lumen. Instead it is either embedding medium (epoxy) or fine clastic mineral with internal reflection. Such examples are frequently observed in Permian coals. Figure 8: is more like exsudatinite (mark the morphology and nature of encroachment on the associated macerals) rather than yellow to orange colour reactive resin (YR) filled along vitrite cracks. Figure 9: is not a lath-shaped resinite (R) body but a blob of clay mineral showing its characteristic internal striations. The body is not in alignment with the surrounding sporinites, which clearly indicates it to be a secondary matter. Figure 10: shows yellow resinite (YR) filled in sporinite. In fact, if properly focussed it will reveal 3–4 sporinites compressed together.

No resinite (resin) as that shown in black and white photographs (Plates 1, 3–5) have been illustrated under fluorescence mode. Calcite is a weakly fluorescing mineral and can be mistaken as resinite. Absence of corpocollinite, cutinite and calcite and low amount of siderite recorded clearly explain the reasons for high amount of resinite erroneously recorded by Saxena *et al.*[11] in the Raniganj coal seams of the Raniganj field.

Conclusions

Utilizing the fluorescence technique, macerals like liptodetrinite, bituminite, exsudatinite, alginite and fluorinite, in addition to fluorescing or perhydrous vitrinite, have been recognized for the first time in Indian Lower Gondwana (Permian) coals of the Raniganj, Singrauli and Rajmahal fields. Thus, these not only add to the number of liptinite macerals but also to their quantity (2–4 times or even more than recorded in normal incident light). This positively indicates that the earlier 'liptinite-poor' concept for Indian Gondwana coals is no longer tenable. High concentration of especially sporinite in some coal bands of Singrauli and Rajmahal fields make them of cannel/semi-cannel coals types.

1. Teichmüller, M., *Int. J. Coal. Geol.*, 1989, **12**, 1–87.
2. Misra, B. K. and Singh, B. D., *Palaeobotanist*, 1994, **42**, 1–13.
3. Teichmüller, M., *Soc. Org. Petrol. Spec. Publ.*, 1982, **1**, 74.
4. Ottenjann, K., *Zeiss Inform.*, 1982, **26–93E**, 40–46.
5. Stach, E., Mackowsky, M.-Th., Teichmüller, M., Taylor, G. H., Chandra, D. and Teichmüller, R., *Stach's Textbook of Coal Petrology*, Gebrüder Borntraeger, Stuttgart, 1982, 3rd ed., p. 535.
6. Misra, B. K., Saxena, R. and Anand-Prakash, *Palaeobotanist*, 1990, **38**, 188–195.
7. Taylor, G. H., Liu, S. Y. and Teichmüller, M., *Int. J. Coal Geol.*, 1991, **18**, 71–85.
8. Teichmüller, M., *Adv. Org. Geochem.*, Pergammon Press, 1986, pp. 581–599.
9. Chakrabarti, N. C., *Ind. Min.*, 1992, **46**, 13–24.
10. Chaudhuri, S. G. and Ghose, S., *Int. J. Coal Geol.*, 1990, **14**, p. 237–253.
11. Saxena, R., Navale, G. K. B., Chandra, D. and Prasad, Y. V. S., *Palaeobotanist*, 1990, **38**, 58–82.
12. Misra, B. K., *Int. J. Coal Geol.*, 1992, **20**, 115–144.

Received 29 September 1993; revised accepted 6 September 1994

Reduced uptake of (^{3}H) thymidine by PHA-responsive human peripheral blood lymphocytes as a function of storage

Awadhesh N. Jha

Centre of Advanced Study in Zoology, Banaras Hindu University, Varanasi 221 005, India
Present address: Brixham Environmental Laboratory, ZENECA Limited, Brixham, Devon TQ5 8BA, UK.

Phytohaemagglutinin (PHA)-responsive lymphocytes from human peripheral blood samples, either irradiated or unirradiated, showed a reduced frequency of (^{3}H) thymidine-labelled cells (detected by autoradiography) as a function of storage duration. The study indicated that storage of blood samples could lead to a delay in the initiation of cell cycle progression and may impart additional source of variability when these cells are used for experimental or clinical purposes.

PERIPHERAL blood lymphocytes (PBL) of human are widely used for clinical and experimental purposes. One of the most widespread uses of PBL is in environmental mutagenesis studies. In such studies, these non-dividing cells are stimulated to divide by mitogens such as phytohaemagglutinin (PHA), under *in vitro* conditions, in order to analyse chromosomal aberrations (CAs) and sister chromatid exchanges (SCEs) from individuals occupationally or accidentally exposed to mutagenic agents. In such studies, CAs and SCEs are analysed respectively from cells which have divided once (first division or M1 cells) or twice (second division or M2 cells) in culture following PHA-stimulation[1]. It is well known that under *in vitro* conditions, a large number of factors influence the cell cycle kinetics of these cells. It is essential to recognize and take into account such factors when biological dosimetry and genetic risk evaluation are made. In this context, it is not always possible to use the blood samples immediately after their collection. A delay, therefore, is inevitable between collection of blood samples and setting up of the cultures especially when they are to be transported from distant places. Only a few reports are available on the effect of storage duration on cell cycle progression of these PHA-responsive cells.

Introduction of fluorescence plus Giemsa (FPG) staining technique[2] has facilitated the identification of cells which have divided once or more[3]. This technique as a prerequisite requires thymidine analogue, 5-bromodeoxyuridine (BrdU), when cells are grown in cultures. Using this technique, increased proportion of M1 cells and subsequent reduction in the frequency of other division cells was shown from the cultures of stored blood samples compared to fresh ones which were irradiated prior to setting up the cultures[4]. In another study on unirradiated blood samples, an increase in the proportion of M1 cells was also noticed as a function of storage duration[5]. It was suggested that when stored blood samples are used for cultures either there is a delay in cell cycle progression[4,5] or a failure in the differential uptake of Giemsa stain by the chromosomes of M2 and subsequent division of cells, leading to misidentification of M2(+) cells as M1 cells[5]. It was also recommended that when blood samples are stored for a day or more, increased amount of BrdU should be added in the cultures for proper differential staining of the chromatids[5]. We, however, found that PHA-responsive lymphocytes, either irradiated or unirradiated, show increased frequency of M1 cells as a function of storage duration. Increasing BrdU concentrations slowed down the cell cycle progression and its higher concentration was not required to elicit proper sister chromatid differential staining when cultures were set up from stored blood samples[6]. The study therefore suggested that storage condition delays the cell cycle progression of these PHA-responsive cells.

In order to check if there is any change in the programme commitment of these cells for the initiation of proliferation due to storage condition which could be responsible for increased proportion of M1 cells, in the present study, an attempt was made to analyse (^{3}H) thymidine uptake by these cells (as a function of storage duration) following mitogen stimulation. Autoradiographic technique was used to analyse (^{3}H) thymidine uptake by lymphocytes from both unirradiated and irradiated blood samples to corroborate our earlier cytogenetic findings[6]. Needless to say, this sensitive technique is being widely used in regulatory and scientific studies to identify mutagenic hazard by measuring unscheduled DNA synthesis (UDS). In these studies, radiolabelling is determined by counting the number of silver grains which develop over the cells[7,8].

Blood samples were collected from two healthy donors using heparnised syringes. Each sample was divided into three equal parts as per storage requirement (0, 1 and 3 days). The fresh blood samples were irradiated with a dose of 2 Gy of X-ray (110 kV 4 mA, @1 Gy min^{-1} at 37°C). Irradiation was performed within 2 h of collection under aseptic conditions at normal atmospheric pressure in 25 ml glass beakers (in the presence of atmospheric oxygen). After 2 h incubation at 37°C, the irradiated blood samples were used for culture. Other two parts of the sample were stored at room temperature (22 ± 2°C) in a biochemical oxygen demand (BOD) incubator and used for cultures after 24 and 72 h (1 and 3 days) respectively. Unirradiated blood samples from both the donors were also used parallel to the irradiated blood samples. Cultures were initiated by adding 0.30 ml of irradiated or unirradiated blood in 5 ml of

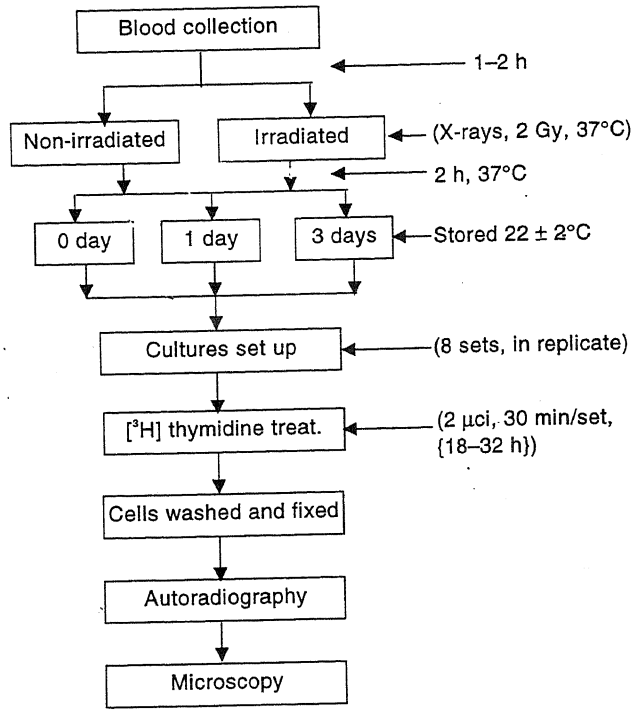


Figure 1. Experimental protocol to study the effect of storage duration on (^{3}H) thymidine uptake by phytohaemagglutinin (PHA)-responsive peripheral blood lymphocytes (PBL).

Eagles minimum essential medium, supplemented with antibiotics, 20% heat-inactivated human AB$^+$ serum and 0.15 ml of PHA (HA 15 Wellcome Reagent Ltd, UK), as described elsewhere in detail[6]. The experimental protocol has been summarized in Figure 1.

As a pilot autoradiographic study, to check the origin time of incorporation of (^{3}H) thymidine in our culture conditions, 12 cultures (in duplicate) were set up from freshly collected blood sample. A 30 min (^{3}H) thymidine pulse (2 µCi ml^{-1}; specific activity 6 Ci Mm^{-1}; BARC, India) was given to these cultures at hourly intervals beginning from 12 h of PHA stimulation up to 24 h. The cells were then fixed in methanol:acetic acid (3:1) followed by very brief washing (2 min) with cold hypotonic (0.56% KCl). Air-dried slides were prepared and processed for autoradiography. The slides were treated with 10% chilled trichloroacetic acid for 5 min, rinsed twice in ethanol and kept under running water for 10 min. After coating with 1:1 dilution of Ilford L4 nuclear emulsion in safe light, the slides were left for exposure in light proof boxes at 4°C for 20 days. The slides were developed in Kodak D 19b developer for 10 min at 10°C and fixed in acid fixer. After drying, they were stained lightly in 2% Giemsa stain and observed under light microscope for the frequency of labelled nuclei. It was found that grains start appearing in cultures 20 h after the PHA stimulation, indicating that initiation of DNA synthesis starts after about 20 h of PHA stimulation in normal lymphocyte cultures.

Table 1. ANOVA for per cent (^{3}H) thymidine-labelled cells from cultures of unirradiated and irradiated blood samples of different individuals after varying hours of PHA stimulation. The cultures were set up on the same day of blood collection or after 1 and 3 days of storage.

Source of variation	Degree of freedom	Sum of squares	Mean of square	Variance ratio	Probability <
Individual	1	13.67	13.67	3.31	ns
Irradiation	1	51.86	51.86	12.56	0.005
Hours after PHA stimulation	6	6762.47	1127.08	272.90	0.005
Storage day	2	268.67	134.34	32.53	0.005
Irradiation × hour	6	31.01	5.17	1.25	ns
Storage day × hour	12	201.53	16.79	4.06	ns
Error	41	169.33	4.13		

ns = not significant.

Therefore, in the subsequent experiments, cultures from both unirradiated and irradiated blood samples after 0, 1 and 3 days of storage were given a 30 min (^{3}H) thymidine pulse at 2 h interval in the same way as mentioned above between 18 and 32 h of PHA stimulation. Eight sets of cultures (in duplicate) for 18–32 h post-PHA-stimulation period, for both irradiated and unirradiated conditions for each storage time and each individual were set up (Figure 1). Slides were prepared, processed for autoradiography as mentioned above, and observed for frequency of labelled nuclei. At least 200 nuclei were scored for each point. Percentage of labelled nuclei was statistically analysed using two-way analysis of variance (ANOVA), as it was considered to be the most appropriate method to check the influence of storage duration on the (^{3}H) thymidine incorporation in the unirradiated and irradiated blood samples used from two different individuals.

The ANOVA for (^{3}H) thymidine-labelled cells in both unirradiated and irradiated samples from two different individuals at different hours of pulse treatment is summarized in Table 1. The two individuals did not show any significant difference for the yield of labelled cells (Figure 2). The results showed that the irradiated samples have significantly less frequency of labelled cells compared to unirradiated ones, in relation to storage duration and advancing hours after PHA-stimulation (Figure 2; Table 1). With the increasing storage duration, independent of advancing hours of pulse treatment, a trend of gradual decrease in the frequency of labelled cells compared to fresh blood samples was also noticed in both unirradiated and irradiated conditions (Figure 2). This difference was found to be significant (Table 1).

Using liquid scintillating analysis, it has been shown that (^{3}H) thymidine incorporation into cellular DNA starts only 16–20 h after PHA-stimulation, and the (^{3}H) thymidine incorporation and activity of DNA polymerase increase linearly as a function of time up to 30 h (refs. 9, 10). This is in broad agreement with the

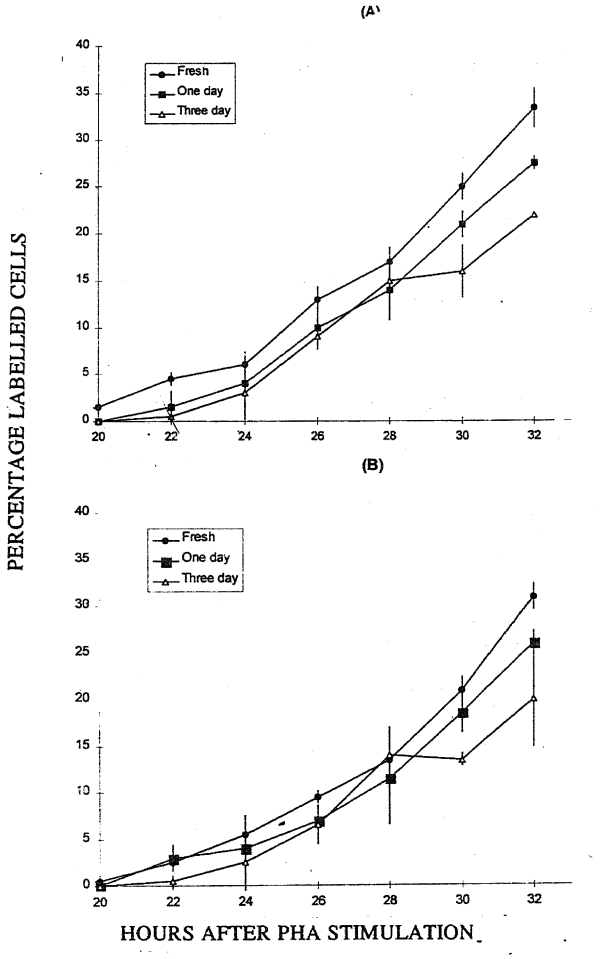


Figure 2. Frequency of (^{3}H) thymidine-labelled lymphocytes in fresh, one day or three day stored unirradiated (*a*) and irradiated (*b*) blood samples after various hours of PHA stimulation.

present autoradiographic study where no significant cellular uptake of (^{3}H) thymidine was observed up to 18 h and then it increased linearly from 20 to 30 h irrespective of storage duration and radiation dose. The pattern for cellular uptake of exogenous deoxyribonucleoside triphosphates (dNTPs) differs in different synchronously or asynchronously dividing cell types and their levels inside the cells are regulated by many factors[11]. Although in the present study, the statistical analysis (ANOVA) did not show any significant inter-individual variation for the frequency of (^{3}H) thymidine-labelled cells, such an inter-individual differences for the response of lymphocytes towards PHA has been a well-documented phenomenon[3]. Significant inter-individual variation for the yield of (^{3}H) thymidine-labelled cells has also been observed when peripheral blood samples from six individuals were labelled between 28 and 34 h following PHA-stimulation[12]. It is therefore likely that independent of the influence of storage condition and radiation treatment, an inter-individual variation would have been observed, if blood samples from more individuals were used in the present study. It will be worth mentioning that in cytogenetic assays using peripheral blood lymphocytes (e.g. micronucleus assay), more variation among the individuals than among repeated experiments on the same individual has been reported[13]. While such inter-individual and flask-to-flask variability for cultured mammalian cells exist, it has been suggested that it has no practical effect on the test outcome[14].

The variability in the (^{3}H) thymidine uptake as a function of storage duration clearly indicates that programmed commitment of cells for the onset of DNA synthesis is altered. This altered commitment for proliferation in turn would influence the cell cycle progression and therefore arrival of the cells at the metaphase stage as observed in our earlier cytogenetic study[6]. A delay in progression from G1 to S, S to G2 or G2 to M phases of the cell cycle may account for an over all effect on cytogenetic dosimetric analysis. In case of X-ray-irradiated lymphocytes, a low dose as 2 cGy has been shown to induce significant variations in cell cycle progression and arrival of the cells at the metaphases[12]. It will be important to mention that there is accumulating evidence indicating that cellular and genetic instability can manifest multiple generations after exposure of cells to agents which either directly or indirectly interact with DNA[15]. It is therefore also possible that the cells might be proportionately delayed in the subsequent cell cycle in the present study. Since the objective of the study was to check whether the appearance of first division metaphases following PHA stimulation as a function of storage is really due to change in the programmed commitment of cells for initiation of proliferation, the cells were fixed at 32 h post-PHA stimulation time. At this time, the first set of metaphases are likely to appear when the cells are processed for chromosomal analysis. With the present experimental approach, it was therefore not possible to verify if there was any delay in the subsequent cell cycle as a function of storage duration.

Since the complexities of cellular growth and division are not precisely clear and products of thousands of genes are required for the normal growth and reproduction of cells[16], it is difficult to assign any specific reason for the delay in cell cycle progression as a function of storage duration. It has been reported that lymphocytes from both pig and human proliferate more slowly in plasma lymphocyte culture than in parallel whole blood culture[17]. Isolated mononuclear leucocytes also exhibit a slower cell cycle progression than whole blood culture[18]. It is well documented that after storage or preservation, changes take place in different components of the blood. These include gradual increase in the osmotic fragility of red blood cells (RBCs), their spontaneous lysis leading to rise in plasma haemoglobin levels besides changes in the nucleotides (including ATP) and phosphorylated carbohydrates[19]. Use of heparin as an

anti-coagulant has also been shown to have no effect on RBC preservation as it lacks dextrose. On storage, its anti-coagulant effect is neutralized by thromboplastic and anti-heparin materials liberated by the cellular elements of the blood[20]. Platelets are considered to be another labile component of blood. It has been suggested that blood stored for more than two days should be considered to be free of viable platelets[21]. The lymphocytes, on the other hand, can be stored as long as four days at room temperature for immunological sub-population analysis[22,23]. These lymphocytes in the absence of optimum blood components like erythrocytes and platelets may progress slowly after PHA-stimulation in culture as observed in plasma lymphocyte or isolated mononuclear leucocyte culture[17,18]. In fact, it has been shown that when rat RBCs are added to purified rat leucocyte cultures, the cell cycle progression is considerably faster. It has been suggested that lymphocytes are released from inhibition to growth and proliferation by the presence of RBCs[24]. In this context, it will be also important to mention that mitogen-stimulated lymphocytes are particularly susceptible to undergo the phenomenon of apoptosis (programmed cell death) either under stressful culture conditions such as serum deprivation or after the induction of damage to cellular DNA[25]. Since storage condition leads to change in the biochemical and physiological properties of the different components of the blood, it is likely that under such stressful condition, the lymphocytes may undergo the phenomenon of apoptosis, resulting into reduced frequency of labelled cells. Parallel to this, the irradiated cells would have probability to undergo DNA damage-induced apoptosis and would show a further reduction in the frequency of (^{3}H) thymidine-labelled cells.

In conclusion, the storage of peripheral blood samples at room temperature delays the cell cycle progression and may impart additional source of variability when used for experimental purposes. It would be important to understand how the interaction of component cells of blood influences the rate of proliferation and the duration of the cell cycle. It would be also interesting to precisely identify different storage labile factors and then by manipulating them, prevent or modify the proliferative response. Such factors, besides other components of the blood may include sub-cellular metabolites, cytosolic and membrane-bound receptors and enzymes.

1. Natarajan, A. T., Tucker, J. D. and Sasaki, M. S., in *Methods to Assess DNA Damage and Repair: Interspecies Comparisons* (eds Tardiff, R. G., Lohman, P. H. M. and Wogan, G. N.), John Wiley & Sons, Chichester, England, 1994, pp. 95–113.
2. Perry, P. and Wolff, S., *Nature*, 1974, **251**, 156–158.
3. Crossen, P. E. and Morgan, W. F., *Exp. Cell Res.*, 1977, **104**, 453–457.
4. Sharma, T. and Das, B. C., *Int. J. Radiat. Biol.*, 1984, **45**, 151–158.
5. Prosser, J. S. and Moquet, J. E., *Experientia*, 1983, **39**, 778–780.
6. Jha, A. N. and Sharma, T., *Indian J. Exp. Biol.*, 1992, **30**, 664–669.
7. Auletta, A. E., Dearfield, K. L. and Cimino, M. C., *Environ. Mol. Mutagenesis.*, 1993, **21**, 38–45.
8. Kennley, J. C., *Mutagenesis*, 1995, **10**, 215–221.
9. Agarwal, S. S. and Loeb, L. A., *Cancer Res.*, 1972, **32**, 107–113.
10. Leonard, J. C. and Merz, T., *Mutation Res.*, 1983, **109**, 111–121.
11. Reichard, P., *Annu. Rev. Biochem.*, 1988, **57**, 349–374.
12. Olivieri, G., Salone, B., Aillaud, M. and Bosi, A., in Third International Symposium on Chromosomal Aberrations, University of Essen, Germany, August 1995 (abstract).
13. Floyd, D. N. and Cassoni, A. M., *Eur. J. Cancer*, 1994, **30**, 615–620.
14. Soper, K. A. and Galloway, S. M., *Mutat. Res.*, 1994, **312**, 139–149.
15. Galloway, S. M., *Environ. Mol. Mutagenesis*, 1994, **23**, 44–53.
16. Lloyd, D., *Biochem. J.*, 1987, **242**, 313–321.
17. Larramendy, M. L., Reigosa, M. A. and Bianchi, M. S., *Mutat. Res.*, 1990, **232**, 63–70.
18. Mehnert, K., During, R., Vogel, W. and Speit, G., *Mutat. Res.*, 1984, **130**, 403–410.
19. Beutler, E., in *Haematology* (eds Williams, W. J., Beutler, E., Erslev, A. J. and Rundles, R. W.), McGraw-Hill Book Company, New York, 1972, pp. 1299.
20. Perkins, H. A., Rolf, M. R. and Acra, D. J., *Transfusion*, 1961, **1**, 151.
21. Mustard, J. F., *Br. J. Haematol.*, 1956, **2**, 17.
22. Shield, C. F., Marlett, P., Smith, A., Gunter, L. and Goldstein, G., *J. Immunol. Methods*, 1983, **62**, 347.
23. Nicholson, J. K. A., Jones, B. M., Cross, G. D. and McDougal, J. S., *J. Immunol. Methods*, 1984, **73**, 29.
24. Wilmer, J. L., Erexon, G. L. and Kligerman, A. D., *Mutation Res.*, 1983, **109**, 231–248.
25. Cohen, J. J., Duke, R. C., Fadok, V. A. and Sellins, K. S., *Annu. Rev. Immunol.*, 1992, **10**, 267–293.

ACKNOWLEDGEMENTS. I am grateful to Professors A. T. Natarajan (Leiden University, The Netherlands) and T. Sharma (Banaras Hindu University, India) for their encouragement and going through the earlier version of this manuscript. This work was supported by IAEA (Vienna) and DAE (Bombay).

Received 21 December 1993; revised accepted 4 November 1995

3.55 Ga old zircon from Singhbhum–Orissa Iron Ore Craton, eastern India

J. N. Goswami, S. Mishra, M. Wiedenbeck†, S. L. Ray* and A. K. Saha*

Physical Research Laboratory, Ahmedabad 380 009, India
†Present address: Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831-6365, USA
*Department of Geology, Presidency College, Calcutta 700 073, India

^{207}Pb/^{206}Pb ages of zircons from the Older Metamorphic Group (OMG), the oldest recognized rock unit from the Singhbhum–Orissa Iron Ore Craton, eastern India, have been determined by ion microprobe.

The zircon ages lie between 3628 and 3196 Ma and cluster around 3.55, 3.4 and 3.2 Ga. The zircons belonging to the first age group occur mostly as subrounded cores or as well rounded to subhedral grains of apparent detrital origin, while the zircons belonging to the younger groups occur as subhedral, prismatic grains or as overgrowths. Our data suggest that the provenance of the sedimentary precursor included granitoid rocks with ages between 3.6 and 3.5 Ga, confirming the presence of Early Archaean crust in the Singhbhum region. The younger age of 3.2 Ga most probably reflects a later metamorphic event which overprinted the OMG rocks. Finally, our data establish an older limit of ~3.5 Ga for the age of sediment deposition.

THE Singhbhum–Orissa Iron Ore Craton of eastern India is a triangular crustal block of Precambrian age covering an area of about 40,000 km^2. The province contains an almost-continuous geological record from the Archaean to the Mid-Proterozoic[1]. The oldest rocks of sedimentary origin recognized in this craton are designated the Older Metamorphic Group (OMG), which include amphibolite facies pelitic schists, arenites, calc-arenites, calc-schists and ortho- and para-amphibolites[1]. The type area of the OMG is around Champua (Figure 1) where it is exposed over an area of about 200 km^2 to the west and southwest of the Champua town. Here the rocks are intruded by biotite–hornblende tonalite and granodiorite gneisses designated the Older Metamorphic Tonalite Group (OMTG). Rock types similar to the OMG (mainly amphibolite grade) also occur as enclaves of varying dimensions throughout the Singhbhum granite batholith[1].

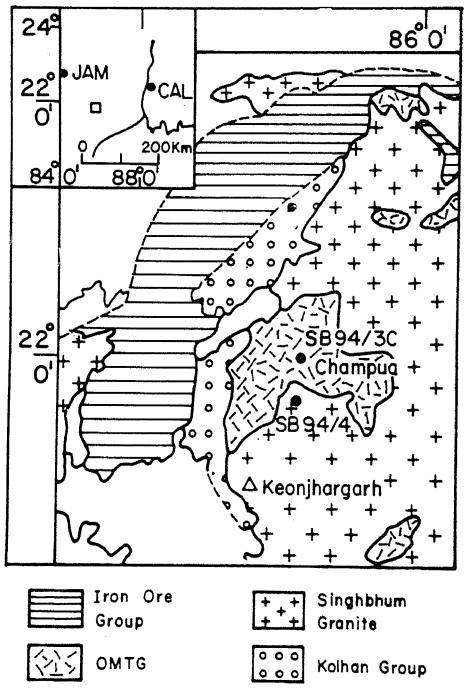


Figure 1. Geological map (after Saha[1]) of the Singhbhum–Orissa Iron Ore Craton around Champua showing the sample locations. The rectangle in the inset indicates the approximate geographical location of the area (JAM – Jamshedpur; CAL – Calcutta).

Geochronologic data from the OMG suite of rocks is restricted to K–Ar and Ar–Ar mineral ages[2,3]. These ages fall in the range of 3.4 to 3.1 Ga and are interpreted as metamorphic ages (see e.g., Saha[1]). This implies that OMG sedimentation took place at or before 3.4 Ga. In this paper we present the results from our ion microprobe study of $^{207}Pb/^{206}Pb$ ages of single zircon grains separated from the OMG metasediments. Our main emphasis is to obtain an age spectrum for the detrital zircon component in the metasediments – the youngest detrital component will provide an older limit for the age of sedimentation. We have considered morphological signatures (e.g. well rounded to subrounded shape with low length to width ratio, resulting from abrasion during sediment transport) to identify zircons of detrital origin (see Figure 2). We have also analysed morphologic overgrowths on single grains, in order to establish the age of metamorphism which has overprinted the OMG.

Two samples of the OMG metasediments were collected for the present study from the type area near Champua (Figure 1). The first sample (SB-94/3C) was collected from a hillock 2 km west of the Champua inspection bungalow. Here the OMG metasediments consist of a sequence of brownish-red muscovite–biotite schists grading to very pure orthoquartzite. The present sample is a quartz-rich member from this sequence. The second sample (SB-94/4) was collected about 15 km away from a site 4 km north of Ramuli village close to the main road connecting Champua to Keonjhargarh. It is an orthoquartzite with mica-rich bands. Although both samples belong to the OMG group, their exact stratigraphic relationship is difficult to establish in the field.

The samples were processed using standard techniques to obtain zircon grains. They were crushed into cm-sized chips and these chips were washed, care being taken to eliminate any material which showed weathered portions. The chips obtained from each sample were then dried in air and were further crushed to < 250 μm using a stainless steel piston and cylinder. The fractions were then washed to remove the fine dust produced during crushing and were dried in air at 60°C. Zircons were concentrated from the washed fraction using an aqueous Na-polytungstate solution (density = 3 g/cc) followed by magnetic separation. Finally, individual zircons were hand-picked using a binocular microscope. The clear and least intensely coloured zircons were mounted on double-sided tape, cast in epoxy and sectioned by polishing. The better grains (based on transparency and simplicity of internal structures) were documented in detail. The grain mounts were thoroughly

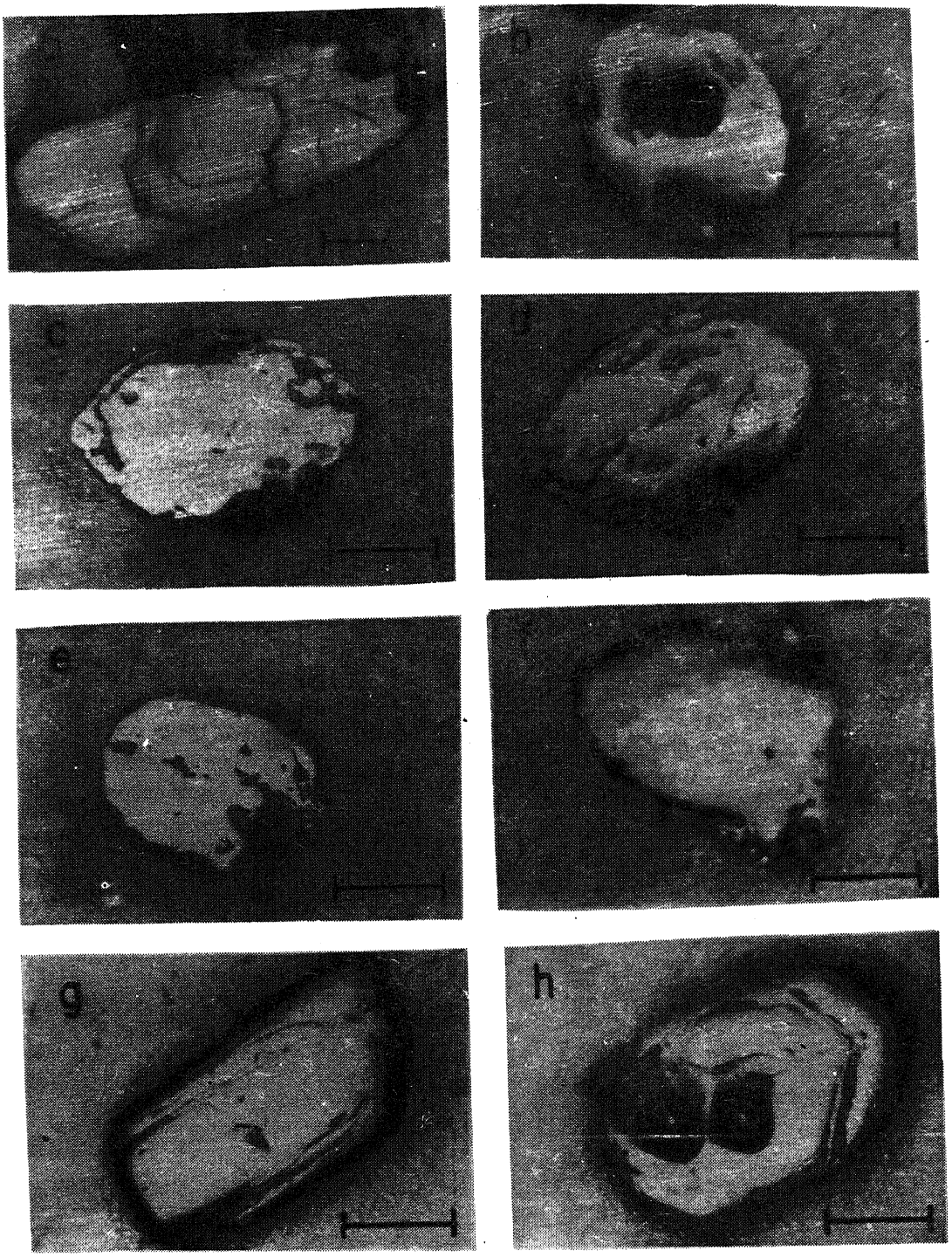


Figure 2. Reflected light photomicrographs of zircon grains recovered from the OMG metasediments. Sputter holes caused by the primary ion beam during Pb-isotopic analyses can be seen in a couple of grains (*b* and *h*). Zircons shown in Figure 2 *a* – *e* belong to the >3.5 Ga group. Scale bar is 50 μm in each case.

cleaned in ethanol using an ultrasonic bath prior to deposition of a 100 nm thick, high-purity gold coat needed for isotopic measurement using the ion microprobe.

Very few good zircons could be extracted from the orthoquartzite (SB-94/3C). The recovered grains are subhedral, prismatic, transparent and appear colourless to brownish in transmitted light. Both homogeneous grains and grains with distinct core-overgrowth morphologies are present. We could recover several good zircons from the sample of the quartzite (SB-94/4). In morphology they are generally homogeneous, though a

Table 1. Ion microprobe Pb-isotopic data for zircons from OMG metasediments, Singhbhum

Analysis no.+	Measured $^{204}Pb/^{206}Pb$	Measured $^{207}Pb/^{206}Pb$	Total ^{206}Pb (counts)	$^{206}Pb^*$ (ppm)	U$ (ppm)	$^{207}Pb/^{206}Pb^*$	Age**
Sample: SB-94/3C							
1A	0.00084	0.3324	19740	65	101	0.3243 ± 69	3591 ± 32
1B	0.00023	0.3188	46170	134	212	0.3165 ± 38	3553 ± 18
1C	0	0.3109	16370	58	92	0.3108 ± 48	3525 ± 24
1D	0.00005	0.2573	78390	265	474	0.2567 ± 27	3227 ± 16
1E	0.00024	0.2609	56505	198	353	0.2583 ± 29	3237 ± 17
2	0.00038	0.2706	45355	124	218	0.2665 ± 39	3286 ± 22
4	0.00149	0.2982	20115	66	112	0.2827 ± 63	3378 ± 34
Sample: SB-94/4							
1A	0.00079	0.2693	56730	152	270	0.2608 ± 37	3251 ± 22
1B	0.00102	0.2630	47100	135	245	0.2518 ± 35	3196 ± 22
2A	0.00040	0.3210	38920	106	168	0.3170 ± 44	3555 ± 21
2B	0.00010	0.3112	26300	100	160	0.3102 ± 40	3522 ± 19
3A	0.00044	0.3105	45835	109	176	0.3061 ± 41	3501 ± 21
3B	0.00141	0.2878	28780	107	185	0.2729 ± 58	3323 ± 33
4	0.00149	0.3063	37295	108	180	0.2910 ± 44	3423 ± 23
7A	0	0.3228	17960	58	90	0.3228 ± 53	3583 ± 25
7B	0.00161	0.3224	20235	60	97	0.3062 ± 60	3502 ± 30
11A	0.00071	0.2956	89430	333	557	0.2883 ± 23	3408 ± 12
11B	0.00010	0.2793	112140	364	622	0.2783 ± 16	3353 ± 9
12	0.00025	0.2948	14410	37	61	0.2923 ± 63	3430 ± 33
13	0.00049	0.3369	12355	18	27	0.3323 ± 83	3628 ± 38

+The alphabets refer to multiple analyses of single zircon.
$Calculated value based on the assumption that the sample has remained a closed system.
*Radiogenic value corrected using 3.5 Ga model common Pb (ref. 8).
**Errors are 1σ.

small fraction have extremely well rounded grain morphologies. These zircons contain both apatite and some unidentified brown prismatic minerals as rare inclusions.

We used a Cameca ims-4f ion microprobe for Pb-isotopic analysis using procedures described earlier[4]. We operate the instrument at high mass resolution (M/ΔM~4600) to resolve all the significant isobaric molecular interferences on the Pb isotope mass spectrum. A 7 nA focused primary beam of $^{16}O^-$ was used to analyse Pb-isotopic composition of ~20 μm domains within individual zircon grains. A single analysis consisted of 15 blocks of data; each block comprised of 5 scans through the mass sequence $^{204}(Zr_2O)$, ^{204}Pb, ^{206}Pb, ^{207}Pb, $^{208}(HfO_2)$ and ^{208}Pb in peak jumping mode. The age of each analysis is based on the weighted mean radiogenic $^{207}Pb/^{206}Pb$ ratio obtained from the data of the 15 quasi-independent blocks. In the absence of U–Pb data, we lack detailed information on the Pb-loss history of the analysed samples. We have, therefore, developed objective criteria to identify Pb-loss as well as multiple age components within our data set. The details of the procedure adopted by us for data acquisition and assessment are reported elsewhere[4–6].

We conducted 10 analyses on three zircon grains from the ortho-quartzite sample SB-94/3C. In three of these analyses the common Pb component was high (as evident from high $^{204}Pb/^{206}Pb$ ratios); these analyses were excluded from the data set (Table 1), thus eliminating those analyses which show obvious evidence for Pb-exchange with the extra-crystalline environment. An unique zircon (grain number 1) recovered from this sample has a distinct subrounded core with a distinct euhedral overgrowth (Figure 2*a*). Three analyses carried out in the core region of this grain (analysis numbers 1A–C) yielded ages ranging between 3591 and 3525 Ma with a mean value of 3551 ± 13 Ma (errors quoted here and rest of the text are 1σ). Two analyses of the overgrowth (numbers 1D, E) yielded a distinctly younger mean age of 3232 ± 12 Ma. Another prismatic zircon having a narrow overgrowth at one end (Figure 2*h*) yielded an age of 3286 ± 22 Ma for its central region. The interior of a third elongated zircon possessing multiple growth zones (Figure 2*g*) has an age of 3378 ± 34 Ma.

Results from 13 analyses conducted on eight zircon grains from the quartzite sample SB-94/4 are also given in Table 1. As in the previous case, the analyses having high $^{204}Pb/^{206}Pb$ ratio are excluded from this data set. One of the zircons analysed from this sample (grain number 7) has an extremely well rounded grain mor-

phology (Figure 2 *b*). Two analyses conducted on this grain (numbers 7A, B) yielded ages of 3583 ± 25 and 3505 ± 30 Ma respectively. Two more zircons – a prismatic grain and the core region of a sub-rounded grain (numbers 2 and 13; see Figures 2 *c*, *d*) – have ages of 3538 (mean of analysis numbers 2A, B) and 3628 Ma respectively (Table 1). Two analyses carried out on either side of a longitudinal fracture in another homogeneous, rounded zircon (grain number 3; Figure 2 *e*) yielded ages of 3501 and 3323 Ma respectively; we suspect that one or both of these ages reflect partial Pb-loss from a single domain which crystallized at or before 3501 Ma. Four analyses conducted on three homogeneous zircons (grain numbers 4, 11 and 12) from this sample (e.g., Figure 2 *f*; grain number 11) yielded intermediate ages around 3.4 Ga. Another homogeneous zircon (grain number 1) from this sample has a distinctly lower mean age of 3223 Ma (analysis number 1A, B). We have also estimated the U content of the zircons based on a close system approximation; these values range from a few tens of ppm to a few hundred ppm (Table 1).

The data shown in Table 1 reveal three groupings: greater than 3.5 Ga, 3.43–3.38 Ga and ~3.2 Ga for the analysed zircons. The age of 3.2 Ga for the overgrowth around the 3.55 Ga old detrital core (SB-94/3C, analysis numbers 1A–F) strongly suggests that the 3.2 Ga age reflects a younger metamorphic event affecting the OMG rocks that led to the growth of zircon on an older core. The morphology of this overgrowth (Figure 2 *a*) does not support a detrital origin for the total core-overgrowth assemblage. We note that the 3.2 Ga signature is seen in zircon grains analysed from both the samples.

Two of the zircons belonging to the >3.5 Ga group are clearly of detrital origin as attested to by their sub-rounded shape (Figures 2 *a*, *b*). Among the other three zircons belonging to this age group, two show distinct evidence of rounding of edges with low length/width ratios (Figures 2 *d*,*e*), while the third one possesses an ambiguous morphology (Figure 2 *c*). We, therefore, conclude that crust of ~3.6 to 3.5 Ga age was a major sialic component within the provenance of the OMG.

The four zircons with ages between 3.38 and 3.43 Ga are all homogeneous, subhedral, prismatic grains and do not show obvious detrital characteristics. Unfortunately, the lack of detailed information on possible Pb-loss history of these grains makes it difficult to attribute these ages to specific geologic events. We note that the lower end of this range is only slightly older than a recently reported Sm–Nd age of 3305 ± 60 Ma (2σ) for the OMG amphibolites[7].

The present study provides several important geochronological constraints on the evolution of the Singhbhum Craton. The zircon ages from the OMG metasediments confirm the presence of an Early Archaean crust in this region. Our data suggest that the provenance of the sedimentary precursor contains a significant granitoid component with ages ~3.6–3.5 Ga. Zircon growth, possibly in response to a regional metamorphic overprinting, has been identified at ~3.2 Ga. Our results also provide an older limit of ~3.5 Ga for the age of sediment deposition.

1. Saha, A. K., *Mem. Geol. Soc. India*, 1994, **27**, 341.
2. Sarkar, S. N., Saha, A. K., Boelrijk, N. A. I. M. and Hebeda, E. H., *Indian J. Earth Sci.*, 1979, **6**, 32–51.
3. Baksi, A. K., Archibald, D. A., Sarkar, S. N. and Saha, A. K., *Can. J. Earth Sci.*, 1987, **24**, 1985–1993.
4. Wiedenbeck, M. and Goswami, J. N., *Geochim. Cosmochim. Acta*, 1994, **58**, 2135–2141.
5. Wiedenbeck, M., Goswami, J. N. and Roy, A. B., *Chem. Geol.*, 1995, in press.
6. Wiedenbeck, M., Goswami, J. N. and Roy, A. B., *J. Geol. Soc. India*, 1995, in press.
7. Sharma, M., Basu, A. R. and Ray, S. L., *Contrib. Mineral. Petrol.*, 1994, **117**, 45–55.
8. Cumming, G. L. and Richards, J. R., *Earth Planet. Sci. Lett.*, 1975, **28**, 155–171.

ACKNOWLEDGEMENT. We thank M. P. Deomurari, V. G. Shah and Ms N. Sinha for technical assistance during various phases of this study.

Received 11 September 1995; accepted 7 October 1995

Significance of the terminal Cretaceous calcareous nannofossil marker *Micula prinsii* at the Cretaceous–Tertiary boundary in the Um Sohryngkew section, Meghalaya, India

Rahul Garg and K. P. Jain

Birbal Sahni Institute of Palaeobotany, 53, University Road, Lucknow 226 007, India

A diagnostic calcareous nannofossil assemblage, representing the terminal Maastrichtian *Micula prinsii* zone, is discovered below the iridium-spike bearing clay-layer in the Um Sohryngkew section, Meghalaya. Its significance for the precise demarcation of the K/T boundary, providing crucial evidence for the presence of the youngest Cretaceous sediments in a continuous, bioturbation-free calcareous shale sequence, is demonstrated.

THE Um Sohryngkew section is located on the western bank of the Um Sohryngkew river, about 1 km northeastwards of the Therria village in the foothills of the Cherrapunji Plateau, Meghalaya, northeastern India

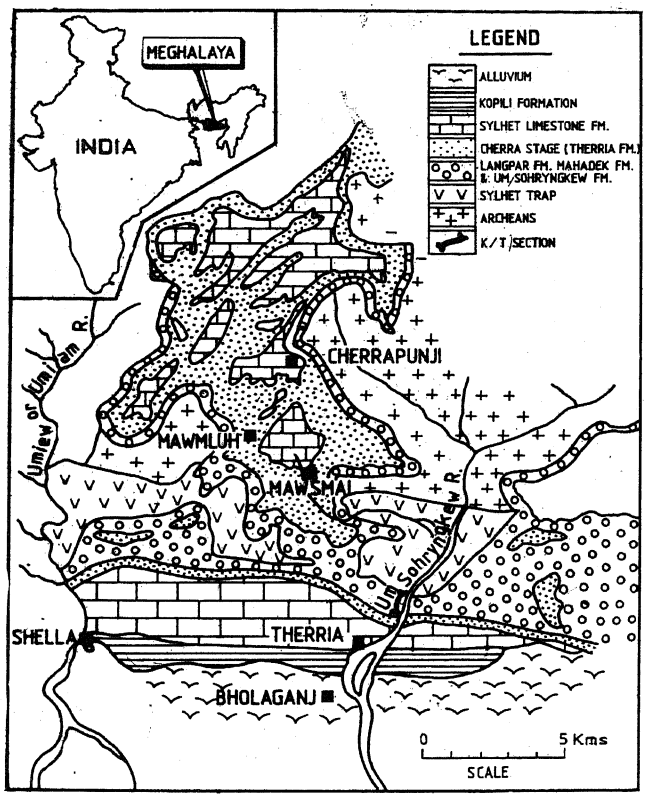


Figure 1. Geological map of the Cherrapunji area, Meghalaya, showing the location of the K/T boundary section (Redrawn after Ghosh, A. M. N., 1940). The 'Cretaceous beds' mapped by Ghosh now include the Um Sohryngkew, Mahadeo and Langpar formations of the Late Crataceous–Paleocene age.

(Figure 1). Cretaceous–Tertiary boundary in this section has been demarcated on the planktonic foraminiferal evidence[1,2]; the top of Maastrichtian has been marked by the abrupt disappearance of a rich *Globotruncana–Heterohelix* suite, while the basal Palaeocene is distinguished by the residual Cretaceous planktonics dominated by *Guembelitria cretacea* (PO Zone) at the base, followed stratigraphically higher by the appearance of a typical Danian *Turborotalia eugubina–T. sabina* assemblage Zone[3]. Geochemical studies across this interval subsequently revealed a significant enrichment of iridium and other siderophyle elements in a 1.5 cm thick clay layer[4,5], coinciding with a major depletion of Cretaceous planktonic foraminifera. Contrary to this, Lahiri *et al.*[6] opined that 'no such widespread break' in planktonic foraminifera at the K/T iridium layer is noticeable, as a Late Maastrichtian assemblage continues stratigraphically higher, suggesting the placement of the boundary about 30 m above this layer. However, these observations have recently been discussed and rejected[7]. Our recent discovery of a rich suite of Danian dinoflagellate cysts and calcareous nannofossils along with the Early Tertiary nautiloid *Hercoglossa kutchensis* Tandon and Srivastava 1980 (Figures 2 *A–C*) below the K/T boundary level suggested by Lahiri *et al.*, also disproves their contention. Jafar and Singh[8] considered the Um Sohryngkew section to be an incomplete K/T boundary sequence due to the absence of the latest Maastrichtian planktonic foraminifer marker *Abathomphalus mayaroensis* and the calcareous nannofossil *Micula prinsii*, and the reported continuation of bored Cretaceous planktonic foraminifera above the iridium layer. They further substantiated this inference by simply considering reports of the presence of the basal Danian planktonic foraminifer markers *Globigerina fringa* and *G. eugubina*[3] to be 'poor' and 'dubious' without providing any substantial arguments.

Thus, the recording of a characteristic latest Maastrichtian calcareous nannofossil assemblage, belonging to the *Micula prinsii* zone, from immediately below the iridium-rich clay layer, confirms not only the existence of the youngest Maastrichtian in this section but also suggests demarcation of K/T boundary below this layer[9,10].

The material for the present study was initially collected during a field trip in February 1987 (in which Jagadish Pandey[3], S. K. Acharyya and M. K. Sen[6] also participated), and subsequently in February 1990, when more closely spaced samples at 5 cm intervals (sample thickness ca. 2–2.5 cm) were collected from a nearly 2 m interval across the clay layer. The studied part of the K/T succession consists of non-bioturbated, thinly laminated calcareous to fine silty shales which are dark grey to ash grey in lower part, followed by a 1.5 cm thick reddish-brown clay layer, overlain by greyish to yellowish-brown calcareous fine silty shales.

These calcareous shales have lately been included within the Mahadeo Formation[1,2]. However, the authors prefer to group these beds within the overlying Langpar Formation, in accordance with the original lithological criteria[11–13]. The Langpar Formation, comprising calcareous shale, carbonaceous shale and argillaceous limestone, conformably overlies the Mahadeo Formation, consisting of glauconitic non-calcareous sandstone and sandy shale, becoming calcareous in the upper part. There is a distinct gradational transition from a greenish-grey, sandstone-shale facies of the Mahadeo Formation into the overlying dark to light grey, calcareous shale-limestone facies of the Langpar Formation. The sequence studied immediately overlies these passage beds.

Abundant and diversified Late Cretaceous calcareous nannofossil assemblages have been recovered from samples underlying the iridium-rich clay layer. They contain the terminal Maastrichtian, low latitude zonal marker *Micula prinsii* Perch-Nielsen 1979 along with *M. murus* (Martini) Bukry 1973. *M. decussata* and *Watznaueria barnesae* are the most abundant and consistent species in the studied interval (Figure 3). This assemblage typically represents the *Micula prinsii* zone of the latest Maastrichtian.

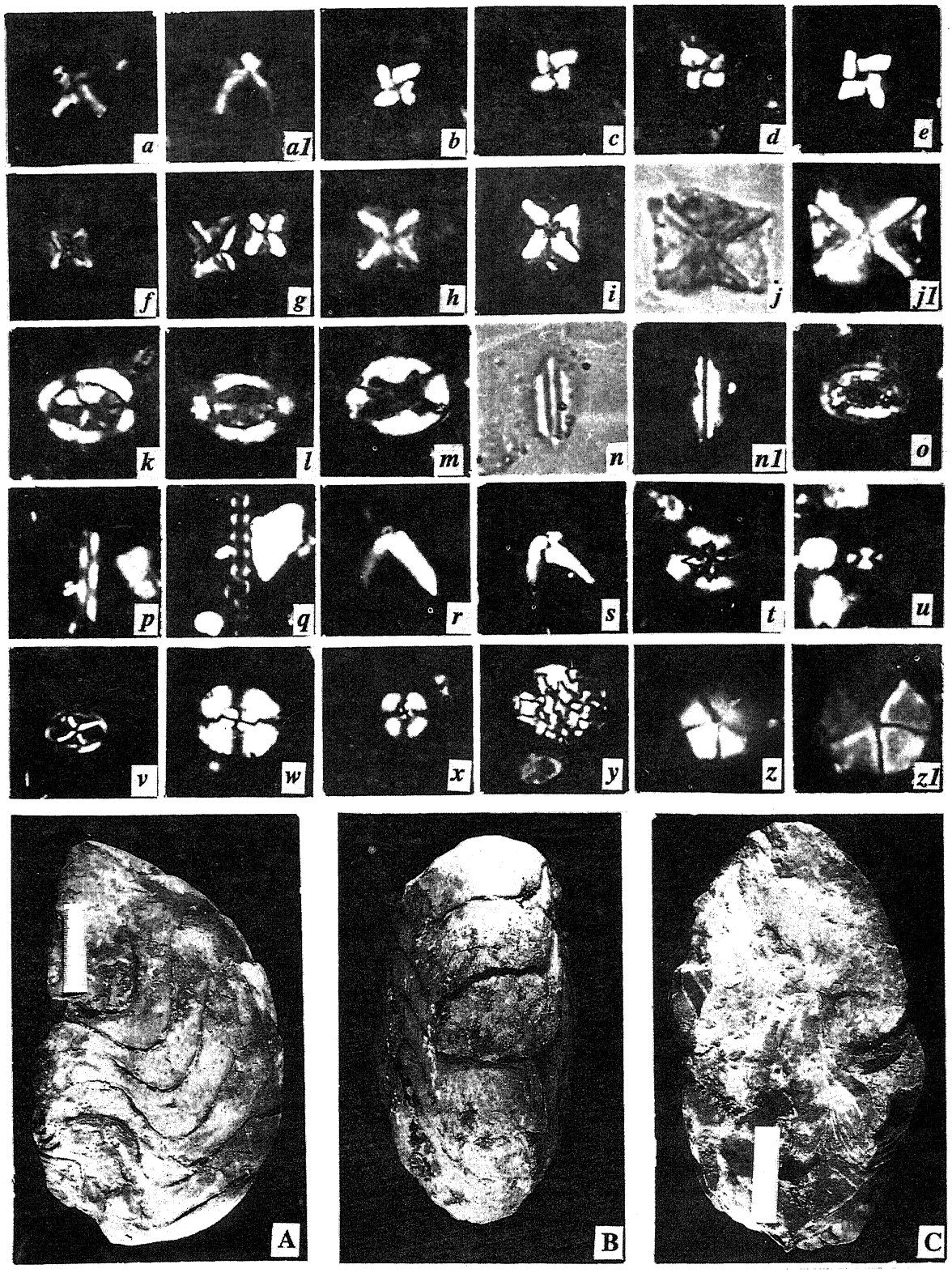


Figure 2. ***a–a1***, *Micula prinsii*; ***b–e***, *Micula murus*; ***f–h***, *Micula decussata*; ***i***, *Micula concava*; ***j, j1***, *Micula* sp.; ***k, l***, *Arkhangelskiella cymbiformis*; ***m***, *Eiffellithus turriseiffellii*; ***n, n1***, *Lithraphidites quadratus*; ***o***, *Cribrosphaerella ehrenbergii*; ***p, q***, *Microrhabdulus decoratus*; ***r, s***, *Ceratolithoides kamptneri*; ***t***, *Cretarhabdus angustiforatus*; ***u***, Gen. et sp. indet.; ***v***, *Chiastozygus amphipons*; ***w***, *Watznaueria barnesae*; ***x***, *Cyclagelosphaera reinhardtii*, ***y***, *Thoracosphaera operculata*; ***z, z1***, *Braarudosphaera bigelowii* (All light micrographs under crossed polarized light except *j* and *n* under single polariser; ×2000). **A–C**, *Hercoglossa kutchensis* (scale = 2 cm).

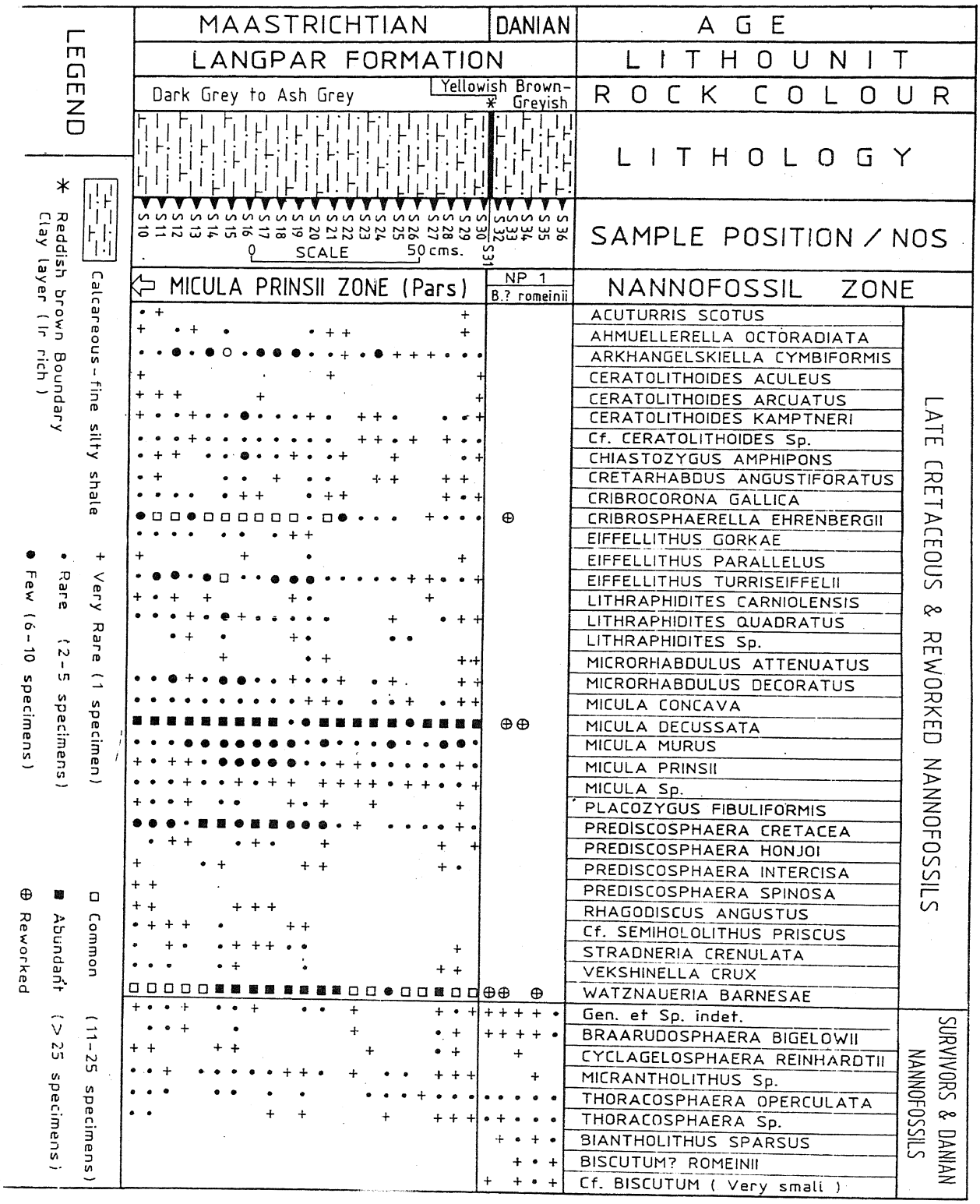


Figure 3. Distribution of Late Maastrichtian calcareous nannofossil species in the Um Sohryngkew section.

The *Micula prinsii* zone was established for the uppermost Maastrichtian[14,15] and can be correlated to the top of the *Micula murus* zone of Thierstein[16] and the *Nephrolithus frequens* (CC26) zone of Sissingh[17]. *M. prinsii*, which evolved from *M. murus*, has its FAD after the first appearance of both *Nephrolithus frequens* and the planktonic foraminifer marker *Abathomphalus mayaroensis*[14,18,19]. It is the most significant species to document the presence of the youngest Cretaceous sequence especially in low and mid latitudes[20,21], e.g. El Kef (Tunisia), Lattengebirge (Germany), Biarritz (France), DSDP site 534 (S. Atlantic)[20]; Barranco del Gredero near Caravaca (Spain)[22] and Hor Harar (Israel)[23].

In Um Sohryngkew section, the rich Maastrichtian nannofossil assemblage abruptly disappears above sample no. S-30 (Figure 3). Nannofossils have not been observed in the iridium-rich clay layer (sample S-31). In the immediately overlying calcareous to fine silty shales, the nannofossil assemblage is greatly reduced containing only rare *Thoracosphaera operculata, Thoracosphaera* sp., *Braarudosphaera bigelowii* and extremely rare *Cyclagelosphaera reinhardtii* and *Micrantholithus* sp. These 'survivor' species belong to a suite of calcareous nannofossils which are believed to have survived the K/T boundary biotic crisis and extend from the Maastrichtian into the basal Danian[20]. Brightly birefringent elliptical to broad elliptical coccoliths (2–5 μm) are found to occur across the K/T boundary, without any appreciable change in size. Rare occurrence of this form is recorded in the Upper Maastrichtian as well as in the basal Danian where it is found to occur consistently in the studied interval. In our opinion, records of this species above the K/T boundary are not due to reworking but seem to represent genuine occurrences. It is documented here as Gen. et sp. indet. and is listed under 'survivor' species for the present (Figure 3).

It is generally believed that a 'continuous' marine K/T boundary section in low latitudes should include the uppermost Cretaceous coccolith zone of *Micula prinsii* followed by the boundary clay within the planktonic foraminiferal zone of *Globigerina fringa*, followed by *G. eugubina*, in the basal Danian[20]. However, the occurrence of a relict survivor Late Cretaceous planktonic foraminiferal faunule belonging to the *Guembelitria cretacea* zone [PO], in the earliest Danian below the *G. fringa* zone is now recognized to be extremely significant in establishing the continuity of a complete K/T boundary sequence[24–27].

The argument of Jafar and Singh[8], that the reported 'continuation' of a relict Cretaceous planktonic foraminiferal assemblage, representing the *G. cretacea* zone, above the iridium-rich layer in the Um Sohryngkew section[3] is abnormal for a complete K/T boundary sequence, is, thus, not valid. Further, the 'absence' of *Micula prinsii* in several Indian sections, including Meghalaya and Andaman Islands postulated by them[8] should be viewed with caution. *M. prinsii* has previously been listed in a 'nannoplankton cocktail' obtained from a solitary mud-volcano sample from Baratnag Island that yielded a mixed assemblage of Campanian–Danian age[28]. However, *M. prinsii* could not be recovered in subsequent studies of mud-volcanoes of the same region[29], thus rendering its earlier report from Andaman Islands to be questionable[30].

According to Pandey[2], the zonal planktonic foraminifer marker, *A. mayaroensis*, is absent in Meghalaya. He consequently instituted the *Globotruncana staurtiformis* zone as a substitute for the Maastrichtian *A. mayaroensis* zone. Jafar and Singh[8] misquoted Bhandari *et al.*[4] by crediting them with noting the occurrence of *A. mayaroensis* in eastern Indian sections, including Meghalaya, but at the same time suggested that this species 'may in fact be missing altogether' from this region. *A. mayaroensis* has been recorded from the Cauvery Basin[31], the Krishna–Godavari Basin[32,33], the Andaman Islands[34,35] and exotic limestone in the Ukhrul Melange Zone of the Nagaland–Manipur Ophiolite belt[36]. Occurrence of a well-preserved and complete assemblage containing *A. mayaroensis* has also been reported from DSDP Site 217 in the Bengal Abyssal plain in the Indian Ocean[37].

In view of the absence of this zonal marker in the Um Sohryngkew section, correlation with the *A. mayaroensis* zone has been suggested based on the occurrence of a rich planktonic foraminiferal suite containing *Globotruncana stuarti, Rugoglobigerina rotundata* and *R. scottii*[2]. However, these taxa do not provide unequivocal evidence for the presence of the youngest Maastrichtian in the section, as none of these is restricted to *A. mayaroensis* zone. Moreover, *A. mayaroensis* is not considered to be a reliable taxon for identifying the K/T boundary, especially in the shallow continental shelf sections as it is geographically and ecologically restricted[27,38]. An alternative biozone, *Pseudotextularia deformis* was proposed for shallow continental shelf sections[26].

We, therefore, conclude that the present discovery of rich calcareous nannofossil assemblages from the *Micula prinsii* zone immediately below the iridium-rich clay layer illustrates the definite presence of terminal Maastrichtian sediments in this section. It apparently substantiates the planktonic foraminiferal record to fulfil the criteria for a biostratigraphically complete and continuous shallow marine KTB sequence. The continuity of the section is also suggested by the occurrence of bioturbation-free calcareous shales, albeit with a modest colour change, across the 1.5 cm thick iridium-rich clay layer which marks the K/T boundary.

1. Pandey, J., Proceedings of VII Colloquium on Indian Micropaleontology and Stratigraphy, Madras (1978) (ed. Rasheed, D. A.), 1981, pp. 70–152.

2. Pandey, J., *J. Palaeontol. Soc. India*, 1981, **25**, 53–74.
3. Pandey, J. and Ravindran, C. N., Proceedings of the Symposium on Palaeocene of India: Limits and Subdivisions (1986), Indian Association of Palynostratigraphers, Lucknow, 1988, vol. 4, pp. 124–184.
4. Bhandari, N., Shukla, P. N. and Pandey, J., *Curr. Sci.*, 1987, **56**, 1003–1005.
5. Bhandari, N., Gupta, M., Pandey, J. and Shukla, P. N., in *Global Catastrophies in Earth History*, The Lunar and Planetary Institute, 1988, pp. 15–16.
6. Lahiri, T. C., Sen, M. K., Raychaudhuri, A. K. and Acharyya, S. K., *Curr. Sci.*, 1988, **57**, 1335–1336.
7. Pandey, J., *Curr. Sci.*, 1990, **59**, 570–575.
8. Jafar, S. A. and Singh, O. P., *Curr. Sci.*, 1992, **62**, 409–413.
9. Jain, K. P., *Curr. Sci.*, 1991, **61**, 697–700.
10. Garg, R. and Jain, K. P., DINO5: Fifth International Conference on Modern and Fossil Dinoflagellates, Zeist, The Netherlands, 1993, p. 43.
11. Medlicott, H. B., *Mem. Geol. Surv. India*, 1869, **7**, 151–207.
12. Ghosh, A. M. N., *Rec. Geol. Surv. India*, 1940, **75**, 1–19.
13. Biswas, B., *Bull. Geol. Min. Met. Soc. India*, 1962, **25**, 1–48.
14. Perch-Nielsen, K., in *Aspekte der Kreide Europas* (ed. Wiedmann, J.), International Union of Geological Sciences, 1979, vol. 6, pp. 223–272.
15. Perch-Nielsen, K., in Cretaceous/Tertiary Boundary Events Symposium, The Maastrichtian and Danian of Denmark (eds Birkelund, T. and Bromley, R. G.), Univ of Copenhagen, 1979, vol. 1, pp. 115–135.
16. Thierstein, H. R., *Mar. Micropaleontol.*, 1976, **1**, 325–362.
17. Sissingh, W., *Geol. Mijnb.*, 1977, **56**, 37–65.
18. Perch-Nielsen, K., in Cretaceous/Tertiary Boundary Events Symposium, Proceedings (eds Christensen, W. K. and Birkelund, T.) Univ of Copenhagen, 1979, vol. 2, pp. 151–155.
19. Perch-Nielsen, K., in *Plankton Stratigraphy* (eds Bolli, H. M., Saunders, J. B. and Perch-Nielsen, K.), Cambridge University Press, 1985, pp. 329–426.
20. Perch-Nielsen, K., McKenzie, J. and He, Q., *Q. Geol. Soc. Am.*, Spl. Pap., 1982, **190**, 353–371.
21. Henriksson, A. S., *Cret. Res.*, 1993, **14**, 59–68.
22. Romein, A. J. T. and Smit, J., in Proceedings of the Koninklijke Nederlandse Akademie van Weltenschappen, Amsterdam, 1981, vol. 84, pp. 295–312.
23. Eshet, Y., Moshkovitz, S., Habib, D., Banjamini, C. and Magaritz, M., *Mar. Micropaleontol.*, 1992, **18**, 199–228.
24. Smit, J., *Geol. Soc. Am.*, 1982, **190**, 329–352.
25. Brinkhuis, H. and Zachariasse, W. J., *Mar. Micropaleontol.*, 1988, **17**, 153–191.
26. Keller, G., *Mar. Micropaleontol.*, 1988, **13**, 239–263.
27. Canudo, J. I., Keller, G. and Molina, E., *Mar. Micropaleontol.*, 1991, **17**, 319–341.
28. Jafar, S. A. *Curr. Sci.*, 1985, **54**, 170–173.
29. Jafar, S. A., Mainali, U. C. and Singh, O. P., *INA Newslett.*, 1989, **11**, 66–68.
30. Jafar, S. A., *N. Jb. Geol. Palaont. Abh.*, 1994, **191**, 251–269.
31. Govindan, A., *Micropaleontol.*, 1972, **18**, 160–193.
32. Govindan, A., in Proceedings of IX Indian Colloquium on Micropalaeontology and Stratigraphy, Udaipur (1981) (ed. Khosla, S. C.) 1982, pp. 81–93.
33. Raju, D. S. N., Ravindran, C. N., Dave, A., Jaiprakash, B. C. and Singh, J., *Geosci. J.*, 1991, **12**, 177–190.
34. Pandey, J., in Proceedings of II Indian Colloquium on Micropalaeontology and Stratigraphy, Lucknow (1972) (ed. Singh, S. N.), 1973, pp. 66–76.
35. Pandey, J. and Rao, V. K., in Proceedings of VI Indian Colloquium on Micropaleontology and Stratigraphy, Varanasi (1976) (ed. Srinivasan, M. S.) 1977, pp. 182–205.
36. Prithiraj, Ch., Mishra, P. K. and Sahni, A., *Curr. Sci.*, 1992, **62**, 478–481.
37. Smit, J. and Romein, A. J. T., *Earth Planet Sci. Lett.*, 1985, **74**, 155–170.
38. Keller, G., *Mar. Micropaleontol.*, 1993, **21**, 1–45.

ACKNOWLEDGEMENTS. We thank Mr Avinesh K. Srivastava for technical help and Mr Pramod K. Bajpai for drafting the figures.

Received 7 July 1993; revised accepted 4 November 1995

Origin and significance of high-grade phosphorite in a sediment core from the continental slope off Goa, India

V. Purnachandra Rao, Ch. M. Rao, M. Thamban, R. Natarajan* and B. R. Rao

National Institute of Oceanography, Dona Paula 403 004, Goa, India
*National Geophysical Research Institute, Uppal Road, Hyderabad 500 007, India

A phosphorite crust was found at 380–390 cm depth interval of a sediment core collected from the topographic high occurring on the continental slope off Goa. This crust is fragile and grey to light brown in colour. Carbonate fluorapatite is the predominant mineral followed by minor pyrite. Thin section studies indicate that it is mostly homogeneous with a few bone fragments and shows porous microstructure. SEM studies show that it contains agglomerated 1–2 μm size apatite globules resembling phosphatized bacteria and coalesced bacteria. Microprobe geochemistry of the phosphorite indicates that it contains 33% P_2O_5 and 50% CaO with other major elements (Si, Al, Ti, Fe, Mg, Na and K) amounting to only 1.8%. The mode of the phosphorite crust formation is discussed in relation to Quaternary phosphorites in upwelling and non-upwelling regions. It is suggested that the initial substrate for the phosphorite crust was most probably a fish coprolite which phosphatized under lower rates of terrigenous sedimentation and calm environmental conditions during the Pleistocene.

IN phosphorites the P_2O_5 content generally exceeds 18% and it may sometimes reach up to 40%. Phosphorites with higher P_2O_5 content are not only of scientific interest but are also of significant economic value. So far, phosphatized limestones consisting of 2–14% P_2O_5 (bulk analyses)[1–5], the sediments[6] with 0.2–2.0% P_2O_5 and glaucony-phosphate facies associated sediments[7] were reported from the western continental margin of India. In this paper, we report the origin and significance of high-grade phosphorites (34% P_2O_5) sporadically occurring at subsurface depths of the sediment cores collected from the topographic high on the continental slope off Goa.

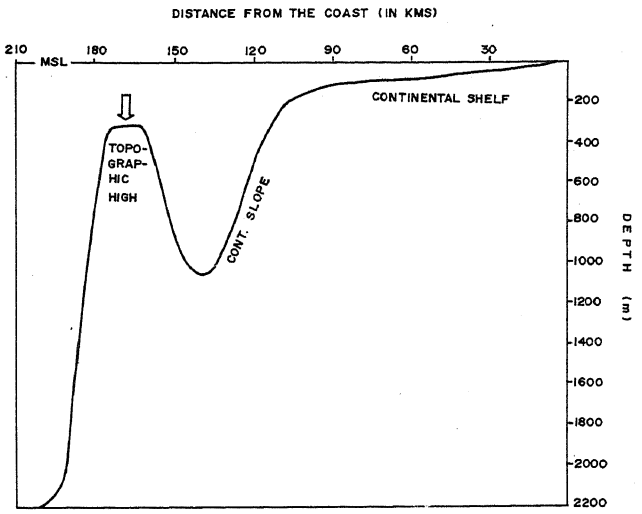


Figure 1. Schematic cross section of the continental margin off Goa showing topographic high on the continental slope; arrow shows the core location (data provided by Dr M. Veerayya).

As a part of Institute's geological exploration programme, more than 30 sediment gravity cores were collected on the western continental margin of India (in the year 1993–94) during different cruises of *O R V Sagar Kanya* and *A A Siderenko*. Most of these cores are from the central and southwestern margin where intense seasonal upwelling has been reported. These cores are distributed at water depths between 20 m and 2000 m and thus are located above, within, and below the oxygen minimum zone (150–1500 m). During subsampling the sediment cores onboard, phosphorite nodules/crusts were found only in three gravity cores collected from the topographic highs occurring on the continental slope off Goa (Figure 1). These phosphorites occur sporadically at one or two sediment intervals at deeper subsurface depths of the cores. All these phosphorites contain about 34% P_2O_5. In the remaining cores there are no phosphorites except a few phosphate grains occurring as traces at different sediment intervals.

The phosphorite crust studied here is from a 4.3 m long gravity core (GC-2) collected on the topographic high in 330 m water depth off Goa (Figure 1) (14°49.6′N, 72°38.8′E) during the VI cruise of *A A Siderenko*. This crust is found at 380–390 cm depth interval of the core. The phosphorite crust is about 3.5 cm long and somewhat flat with thickness of about 1.5 cm (Figure 2 *a*). It is grey on one side and light brown on the other. It is friable and similar to F (friable)-phosphates reported by Garrison and Kastner[8]. Both the surfaces show grooves (Figure 2 *a*) and small boring cavities, indicating that it was indurated under hardground conditions. Bulk mineralogy by X-ray diffraction (Figure 3) shows that the carbonate fluorapatite is the only mineral phase. Pyrite, however, occurs in acid insoluble residue. The bulk chemistry analysed by ICP-AES shows that the crust consists of about 34% P_2O_5 and 49% CaO. The major element geochemistry of the phosphorite, analysed by microprobe, significantly differs from that of the Modern/Quaternary phosphorite nodules/crusts in other upwelling/non-upwelling regions (see Table 1) and consists of more P and Ca and very low content of other major elements. Bone fragments contain relatively less P_2O_5 than the matrix (Table 1).

Thin section of the phosphorite indicates that it is mostly homogeneous with a few scattered bone fragments and fish remains and shows porous structure (Figures 2 *b*, *c*). Pyrite occurs as small black grains. The bones are susceptible to degradative transformation and show different stages of phosphomicritization (Figures 2 *d*, *e*). Some bones are partly micritized (Figure 2 *e*) and others are completely micritized and transformed into undifferentiated peloids (Figure 2 *f*). Elsewhere, the bones are surrounded by black material of possible organic origin. SEM observations on freshly broken surfaces of the phosphorite attest the thin section findings. It consists of abundant porous microstructure with adjacent moderately compacted areas (Figure 4 *a*). Within the porous area, apatite everywhere occurs as agglomerated globules of different size (0.5–2 μm) (Figure 4 *b*). These globules resemble phosphatized bacteria and coalesced bacterial structures[9,10]. The compact areas seem to correspond to thick coalescence of globules and globule aggregates. Figure 4 *c* shows the pyrite aggregate composed of many cubes and octoids of pyrite and porous apatite structure corresponding to globules.

The sediments of the core are clayey at the surface and up to 40 cm depth of the core and clayey sand/silty sands occur below. The organic carbon content in the upper 40 cm of the core is high (about 8%) and in the remaining core, it ranges from 0.8 to 3.7%, with an average value of 1.0%. In contrast, the calcium carbonate content is low (up to 43%) in the upper 40 cm and increases down core with values ranging from 60 to 80%. Coarse fraction studies indicate that in the upper part of the core, planktonic foraminifers show solution features and contain several keels of planktonic foraminifers and pyrite encrusted/infilled foraminifers. In the remaining part of the core, the coarse fraction consists of abundant planktonic forams, pteropods, shell fragments, molluscs and otoliths distributed more or less uniformly in the core. Corals encrusted on a small gastropod shell (Figure 5) were found at a sediment interval 360–370 cm), close to that of phophorite depth interval (380–390 cm).

Some of the Modern phosphorites (off Peru and SW Africa) occur in upwelling regions[11,12], while others (Eastern Australia phosphorites) are in weak/non-upwelling regions[13]. As the western margin of India is concerned, (a) intense seasonal upwelling occurs on the southwestern continental margin of India, (b) organic carbon values range from 3–16% in the surficial sediments of the continental slope[14] and up to 8% in the

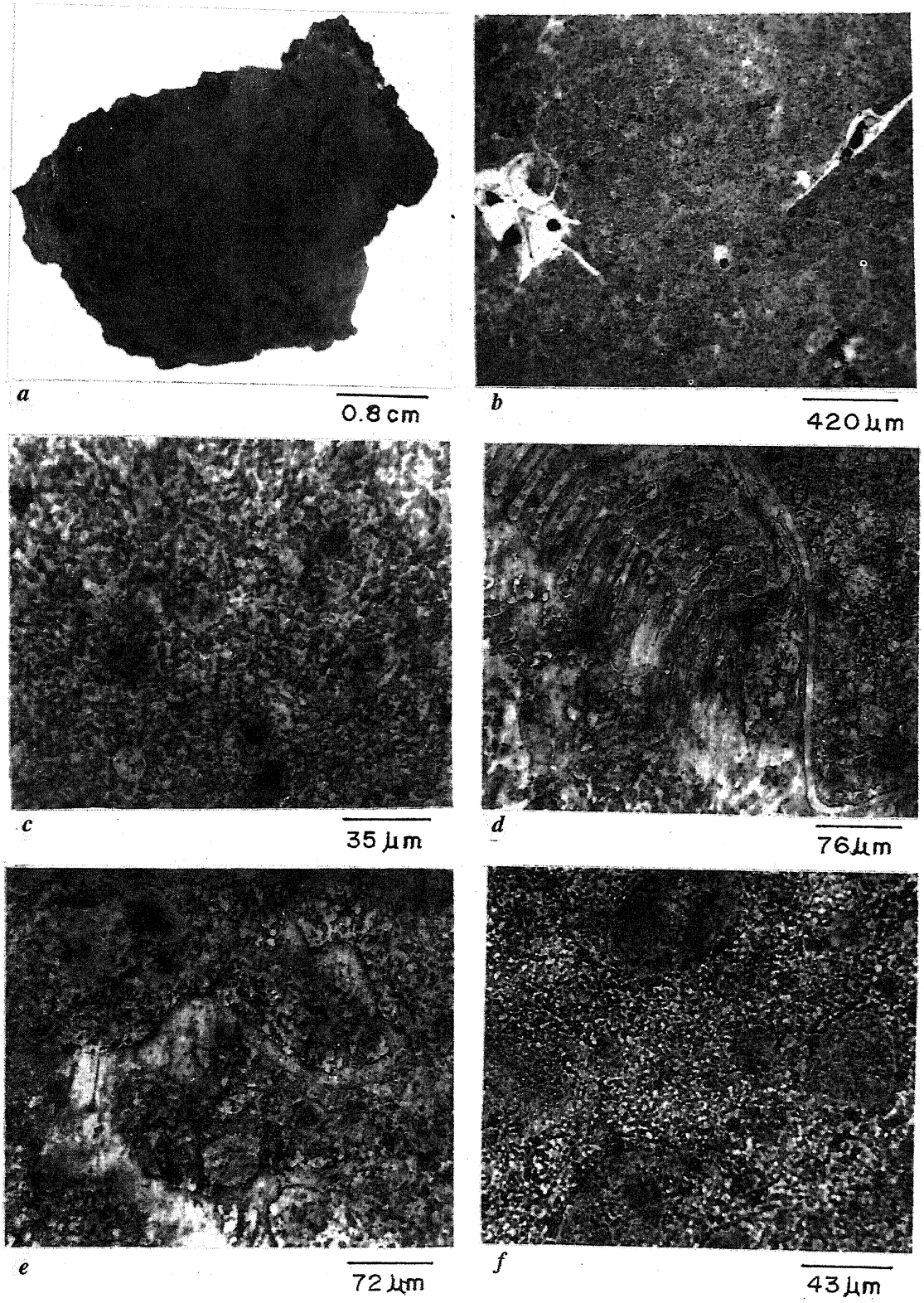


Figure 2. *a*, Phosphorite crust; (*b–f*) are thin section photomicrographs; *b*, at low magnification showing homogeneous apatite structure with a few bone fragments; *c*, high magnification showing porous microstructure; *d*, fish bones being micritized; *e*, bone fragment showing phosphomicritization; *f*, the end product of phosphatization of bones showing undifferentiated phosphate peloids.

Table 1. Geochemistry of the phosphorite and their comparison with modern phosphorite nodules from other regions

	Phosphorite crust from the Indian margin (upwelling region)		Modern phosphorites from		
			other upwelling regions		weak/non-upwelling region
Element	1 (15)*	2 (3)	3 (21)	4 (46)	5 (4)
SiO_2	0.11	0.06	8.79	19.39	24.00
Al_2O_3	0.33	–	2.68	1.95	3.53
TiO_2	0.03	0.03	ND	ND	ND
FeO	0.39	0.13	1.28	6.95	4.50
MnO	0.02	0.01	ND	ND	ND
MgO	0.81	0.52	0.99	1.41	1.31
CaO	50.17	49.00	42.00	37.28	31.90
Na_2O	–	–	1.08	0.90	0.90
K_2O	0.10	0.12	ND	1.38	0.89
P_2O_5	33.26	30.72	27.58	17.42	11.00

* Average of number of analyses, ND – not determined,
1 & 2 This study, microprobe analyses of phosphorite crust from the western margin of India; 1. analyses of matrix; 2. analyses of bone fragments.
3 Phosphorite nodules from Peru margin (microprobe analyses)[11].
4 Phosphorite nodules from Namibia (microprobe analyses)[12].
5 Eastern Australian margin non-ferruginous phosphorite nodule[26].

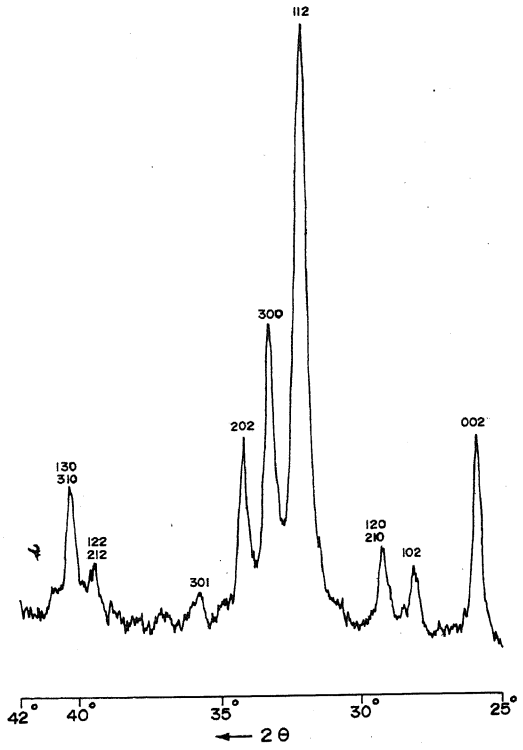


Figure 3. X-ray diffractogram of the phosphorite (bulk analysis), showing all the reflections are due to carbonate fluorapatite.

upper 1 m sediments of the gravity cores (M. Thamban, unpublished work) and are comparable to the organic carbon values of the Peru and SW African margins and (c) oxygen minimum zone (150–1500 m) intersects the continental slope. In spite of all ((a)–(c)) these upwelling-induced conditions, phosphorites have not been observed in the upper one meter sediments of all the thirty gravity cores. We, therefore, suggest that just having upwelling is not sufficient enough to produce phosphorites. Glenn *et al.*[15] recently reviewed the environmental conditions for phosphorite formation and suggested the influence of several factors on phosphogenesis. As the sediments in the cores are carbonate-rich, carbonate ion concentrations in the pore waters[16] and rate of sedimentation are probably major controlling and inhibiting factors for the formation of modern phosphorite in this margin.

Modern/Quaternary phosphorites mostly occur in the form of nodules/crusts. The initial substrate (host) for phosphorite is part of sediment and obviously the phosphorite nodule/crust is heterogeneous and contains partly replaced and/or unaltered sediment constituents such as diatoms, foraminifers, quartz, clay minerals, etc. The geochemistry of this type of diagenetic nodules/crusts is shown by higher contents of other major elements along with higher P values (see Table 1) supporting the existence of sediment constituents. Although the calcareous skeletal dominated sediments are associated with the phosphorite under investigation, it neither contains carbonate skeletals nor other sediment constituents except bone fragments (Figure 2). Geochemistry also shows higher P and Ca values and very low content of other major elements than the modern phosphorite nodules/crusts (see Table 1). These suggest that the host substrate for phosphorite may not be the associated sediment. Similarly, based on petrography, we suggest that the host substrate in our phosphorite is not identical to those in pelletal phosphorites[17] and phosphatized microstromatolites[18].

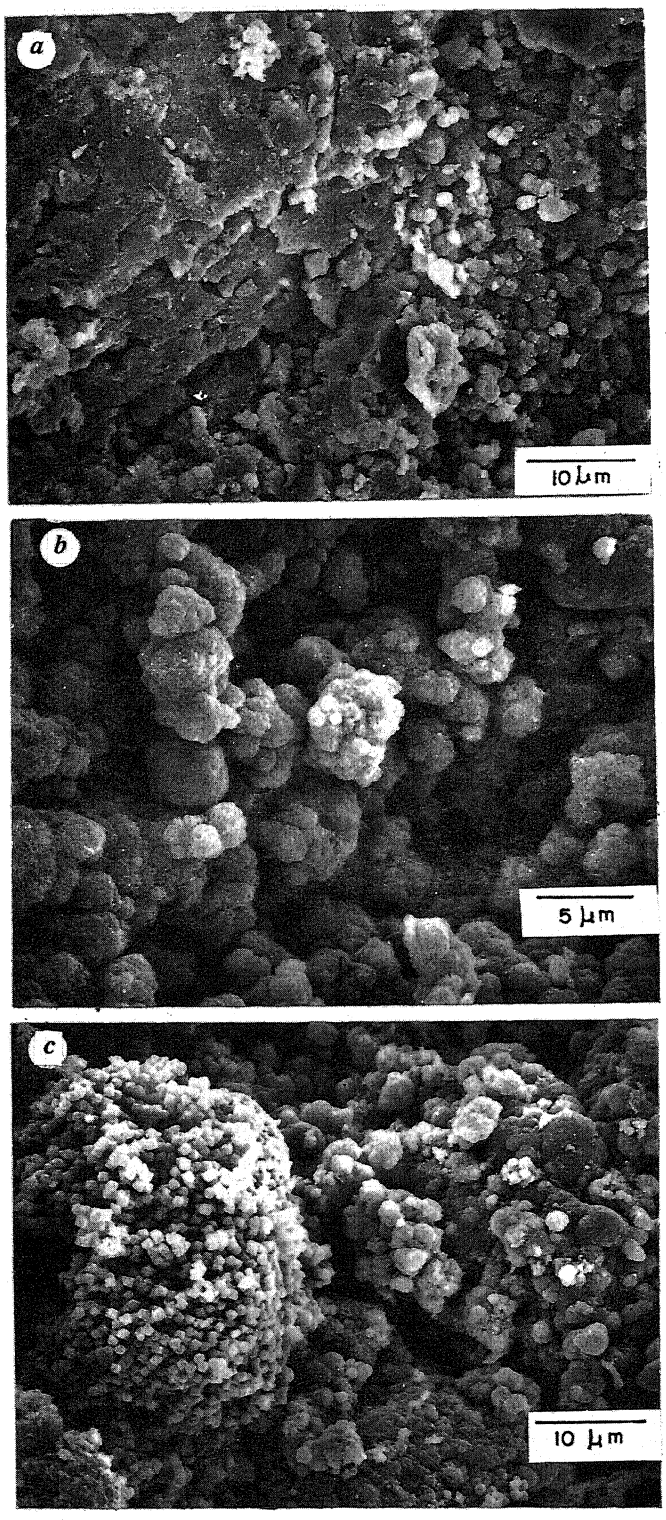


Figure 4. *a–c*, SEM photographs. *a*, apatite microstructure with moderately compact area with adjacent porous areas. *b*, porous area showing agglomerated apatite globules resembling phosphatized bacteria. *c*, Pyrite aggregate showing cubes and octoids of pyrite with adjacent area showing porous apatite microstructure corresponding to globules.

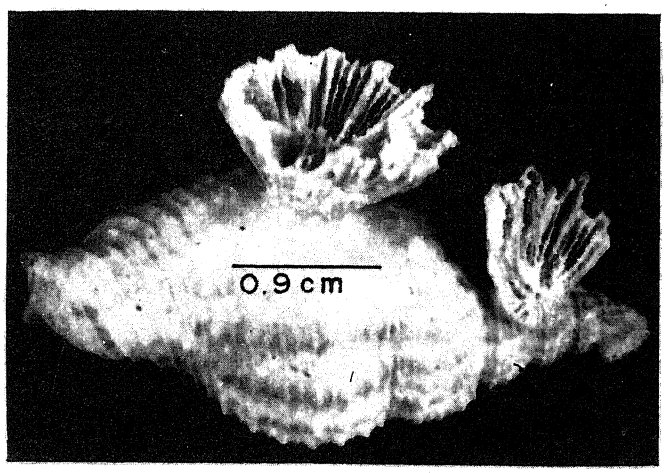


Figure 5. Photograph showing corals encrusted on a small gastropod shell.

The only possibility is that the initial substrate for the phosphorite crust may be a coprolite. Cayeux[19] described coprolites which show typical homogeneous and porous microstructure with few fish remains. More recently, Lamboy[20] reported copronodules from ODP cores of the Peru margin. Subsequent studies[9,10] have documented the homogeneous nannostructure of fish coprolites from various phosphorites including those from the ODP cores of the Oman margin. The observations in our phosphorite are in good agreement with the reports of these workers. It has been suggested that the faeces of carnivores contain highest P_2O_5 content (up to 11%) as they contain undigested remains of bone material and, those of herbivores seem to contain not more than 0.2–0.4% P_2O_5 (cf Slansky[21]). We therefore suggest that the initial excrement may belong to carnivore fishes and become phosphatized. The phosphorite thus produced subsequently became flat due to the overburden of sediment. Absence of hydroxyl apatite reflections in X-ray diffractogrammes, in spite of the existence of bone fragments and fish remains, may be due to the fact that the replacement of hydroxyl ion of the bone by fluoride ion is an ion exchange process and does not require major chemical alteration[21] and/or the total hydroxyl apatite in the sample is <5%.

The precise mechanism of apatite precipitation in coprolites is not known but may be similar to the one reported for authigenic minerals. (a) The major sources for phosphate were most probably the P released during early diagenetic degradation of organic matter and bone fragments associated with the initial excrement. (b) It is also known that dense populations of microbes inhabit the intestines of living vertebrates and their faeces normally contain abundant bacteria[22]. The presence of only agglomerated microspheres (Figure 4 *b*) resembling

phosphatized bacteria everywhere in the phosphorite may indicate the confined environmental conditions favouring the growth of only one type of microbial group. (c) The associated sediments do not contain pyrite whereas, the phosphorite crust contains several pyrite grains implying that the initial excrement itself may have acted as a restricted suboxic/anoxic microenvironment. In view of (a)–(c), we suggest that the microbes existing within the initial excrement assimilated P and F from within and also from the overlying waters, precipitated carbonate fluorapatite and phosphatized only coprolite in the confined microenvironmental conditions.

The initial excrements of fish are very fragile and may rapidly disintegrate with any small disturbance. The initial excrement was survived, phosphatized and indurated into a crust implying that very calm environmental conditions must have existed at the site during phosphatization. Growth of corals on a small gastropod shell (Figure 5) occurs close to the sediment interval of phosphorite crust. The transportation of these corals from elsewhere is ruled out as the sediment core under study was taken from a topographic high, an isolated feature on the continental slope (Figure 1). The encrustation of coral even on a small gastropod shell implies the prevalence of quiet conditions and low rates of terrigenous sedimentation. These conditions thus corroborate the phosphatization conditions of coprolite.

As the phosphorite crust is not occurring in upper 50 cm of the core, it is not a modern phosphorite. The age of the sediments in the core depends on the rates of sedimentation. The rate of sedimentation on the continental slope is about 4 mm/yr (ref. 23); these rates, however, cannot be used for our core as it was taken from a topographic high. Colour and texture of the sediments, organic carbon and calcium carbonate contents of the sediments significantly differ and vary in the sediments of the upper 40 cm of the core and the sediments below. This depth representing changing sedimentary conditions may correspond to Holocene/Pleistocene boundary. The sediment cores collected in the vicinity (Laccadive Sea) also showed similar sedimentary change at about 40 cm depth interval and based on the radiocarbon dates this depth interval has been suggested as Pleistocene/Holocene transition[24]. The rate of sedimentation during the Pleistocene was apparently low as evidenced by the phosphorite and coral encrusted gastropods at different levels. Even if we apply average sedimentation rate (2.2 cm/10^3 yr) recorded in the adjacent Arabian Basin[25], the sediments at 380–390 cm (phosphorite) interval might have deposited during the upper Pleistocene.

1. Baturin, G. N., *Developments in Sedimentology*, Elsevier, Amsterdam, 1981, pp. 343.
2. Nair, R. R., *Bull. Natl. Inst. Oceanogr. India*, 1985, **18**, 273–279.
3. Borole, D. V., Rajagopalan, G. and Somayajulu, B. L. K., *Mar. Geol.*, 1987, **78**, 161–165.
4. Rao, V. P. and Nair, R. R., *Mar. Geol.*, 1988, **84**, 105–110.
5. Rao, V. P., Natarajan, R., Parthiban, G. and Mascarenhas, A., *Mar. Geol.*, 1990, **95**, 17–29.
6. Rao, Ch. M., Paropakari, A. L., Mascarenhas, A. and Murty, P. S. N., *J. Geol. Soc. India*, 1987, **30**, 423–438.
7. Rao, V. P., Lamboy, M. and Dupeuble, P. A., *Mar. Geol.*, 1993, **111**, 133–158.
8. Garrision, R. E. and Kastner, M., Proceedings ODP Science Results, 1990, vol. 112, pp. 111–134.
9. Lamboy, M., Rao, V. P., Ahmed, E. and Azzouzi, N., *Mar. Geol.*, 1994, **20**, 373–383.
10. Rao, V. P. and Lamboy, M., *Ocean. Acta.*, 1995, in press.
11. Burnett, W. C., *Geol. Soc. Am. Bull.*, 1977, **88**, 813–823.
12. Bremner, J. M. and Rogers, J., in *Phosphorite Deposits of the World. Neogene to Modern Phosphorites* (eds Burnett, W. C. and Riggs, S. R.), Cambridge University Press, 1989, vol. 3, pp. 143–152.
13. O'Brien, G. W., Milnes, A. R., Veeh, H. H., Heggie, D. T., Riggs, S. R., Cullen, D. J., Marshall, J. F. and Cook, P. J., in *Phosphorite Research and Development* (eds Notholt, A. J. G. and Jarvis, I.), Spec. Publication Geol. Soc. London, 1990, vol. 52, pp. 61–86.
14. Paropkari, A. L., Rao, Ch. M. and Murty, P. S. N., in *Petroleum Geochemistry and Exploration in the Afro-Asian Regions* (eds Kumar, R. K., Dwivedi, P., Banerjee, B. and Guptha, V.), Balkeme Rotterdam, 1985, pp. 347–360.
15. Glenn, C. R., Follmi, K. B., Riggs, S. R. *et al. Eclogae Geol. Helv.*, 1994, **87**, 747–788.
16. Glenn, C. R., in *Phosphate Deposits of the World, Neogene and Modern Phosphorites* (eds Burnett, W. C. and Riggs, S. R.), Cambridge University Press, Cambridge, 1990, pp. 46–61.
17. Cook, P. J. and McElhinny, M. W., *Econ. Geol.*, 1979, **74**, 314–330.
18. Soudry, D., *Sediment. Geol.*, 1992, **80**, 77–88.
19. Cayeux, L., *Les phosphates de chaux sedimentaires de France. III Etude de gites mineraux de la France*, Imprimerie Nationale, Paris, 1950.
20. Lamboy, M., *Mar. Geol.*, 1994, **118**, 5–22.
21. Slansky, M., *Studies in Geology*, North Oxford, Oxford and B.R.G.M. Orleans, 1986, pp. 210.
22. Thulborn, R. A., *Paleogeogr. Paleoclimatol. Palaeoecol.*, 1991, **83**, 341–366.
23. Borole, D. V., *Mar. Geol.*, 1988, **82**, 285–291.
24. Sankar, R. and Manjunatha, B., *J. Geol. Soc. India*, 1995, **45**, 689–694.
25. Sarkar, A., Ramesh, R., Bhattacharya, S. K. and Rajagopalan, G., *Nature*, 1990, **343**, 549–551.
26. Cook, P. J. and O'Brien, G. W., in *Phosphate Deposits of the World, Neogene and Modern Phosphorites*, (eds Burnett, W. C. and Riggs, S. R.), Cambridge University Press, 1989, vol. 3, pp. 98–115.

ACKNOWLEDGEMENTS. We thank Dr E. Desa and Shri R. R. Nair for their encouragement. Core locations on continental margins for *A A Siderenko-VI* cruise were finalized after discussions with Dr M. Veerayya and based on bathymetric and subsurface features of the SW continental margin. We thank Dr M. Veerayya for this help. Dr M. Sudhakar and other participants helped in collection of the cores. We thank Prof. Michel Lamboy, University of Rouen, France for SEM photographs and Dr V. K. Banakar for chemical analyses with ICP-AES. One of us (MT) thanks CSIR for a Fellowship. We thank U. Sirsat for photography. This paper is a contribution to IGCP project 325-Correlation of Paleogeography with phosphorites and other authigenic minerals.

Received 6 July 1995; revised accepted 4 November 1995

Palynofloral analysis of sections from Bharathi and Kundara clay mines of Kerala Basin: Palaeoecological and tectonic perspective

K. P. N. Kumaran, K. Soman*, C. V. Kamble and Ancy Joseph*

Geology & Palaeontology Group, Agharkar Research Institute, G. G. Agarkar Road, Pune 411 004, India
*Centre for Earth Science Studies, Trivandrum 695 031, India

The carbonaceous clay beds and lenses of peat belonging to Bharathi and Kundara clay mine sections of Warkalli Formation have yielded a well-preserved palynoflora. The assemblage represents a mixture of ecological groups such as low-land, freshwater, and sandy beach elements. The recognition of salt glands signifies that the palynoflora has affinity with mangrove swamp/coastal vegetation. The relative abundance of fungal remains is indicative of a high organic input environment and a warm, humid climate with heavy rainfall during the deposition. Occurrence of *Compositoipollenites, Malvacearumpollis, Quilonipollenites, Trilatiporites* and *Ctenolophonidites costatus* suggests Neogene (Miocene) age. The assemblage is contemporaneous of 'Ratnagiri Beds' of Maharashtra coast as they were deposited in a near-shore environment and their floristic composition is comparable. Nevertheless, the Warkalli deposits remained relatively stable unlike the 'Ratnagiri Beds' which have been subjected to post-depositional displacement caused by tectonic activity. Analysis of accrued fossil pollen evidence and disjunct distribution of plant taxa in a tectonic perspective have also been discussed.

DURING our studies on the kaolin deposits in the coastal tract of South Kerala, carbonaceous clay beds and small bands and lenses of peat have been observed in the ceramic clay mine sections of Bharathi and Kundara[1]. As these deposits contain rich organic matter, the samples have been subjected to palynological study in order to ascertain the floristic composition and also to derive palaeoecological information, which is the main objective of this communication. The samples investigated belong to the Upper Tertiary sedimentary sequence (Warkalli Formation) of the Kerala Basin[2].

The details of location (Figure 1) and clay mine profiles from where the samples were obtained are given below:

The Bharathi clay mine (8°37′38″:76°49′1″) is located at Sasthavattom of Thiruvananthapuram district. The mine section (Figure 2 *a*) shows 10–13 m thick bed of white clay, overlain by a soil column containing laterite pebbles, of thickness up to 6 m. A lens of peat (CC 28)

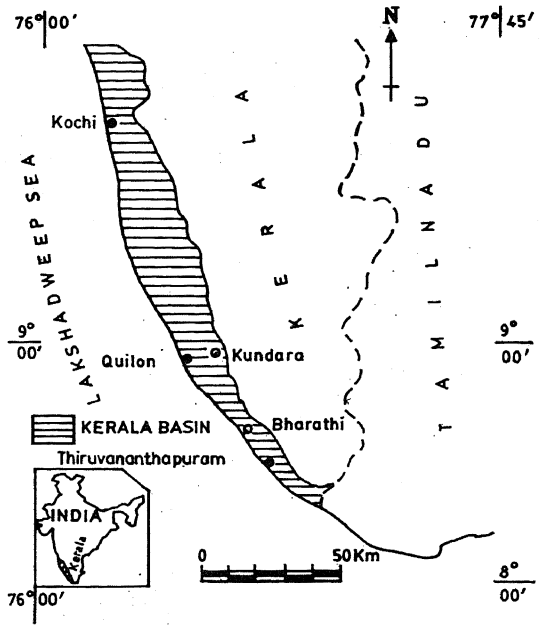


Figure 1. Map showing sampling localities.

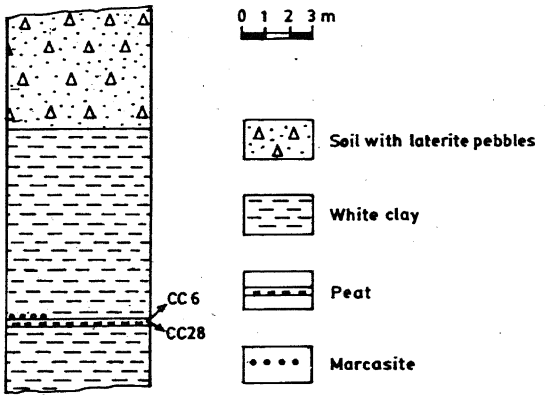


Figure 2 *a*. Synoptic profile of Bharathi clay mine, Sasthavattom.

with associated marcasite occurs at a depth of 13.5 m from the surface within the white clay bed.

The Kundara clay mine section (8°58′:76°40′6″) shows a massive thick sequence unconformably overlying the Precambrian crystalline gneiss (Figure 2 *b*). The maximum height of the southern wall of the mine is 50.7 m and the basal portion consists of residual clay having a thickness of 8 m. Residual clay is overlain by a sequence of sedimentary clay having thickness of 40 m and within this, carbonaceous clays, ball clays, variegated clays and pink clays occur alternatively. Although both the sections are separated by a considerable distance, the lithological features almost remain the same. Further field observations and details of synoptic profiles of Bharathi and Kundara clay mines have been dealt separately in the previous work[1,3].

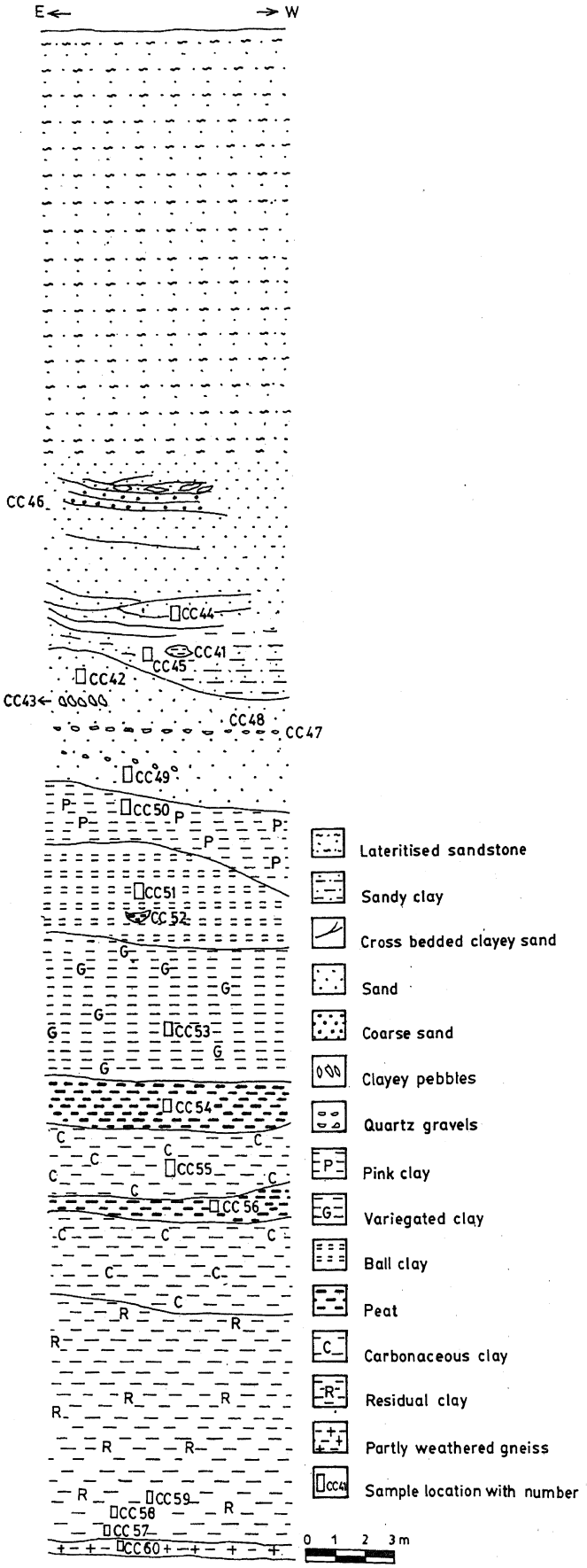


Figure 2 *b*. Synoptic profile of Kundara clay mine.

Table 1. Details of samples and location

Sample	Lithology	Location
CC-6	Carbonized wood	Bharathy Clay Mine, Sasthavattom
CC-28	Peat	Bharathy Clay Mine, Sasthavattom
CC-41	White clay	Kerala Ceramics Clay Mine, Kundara
CC-45	White clay	Kerala Ceramics Clay Mine, Kundara
CC-50	Pink clay	Kerala Ceramics Clay Mine, Kundara
CC-51	Ball clay	Kerala Ceramics Clay Mine, Kundara
CC-52	Peat	Kerala Ceramics Clay Mine, Kundara
CC-53	Variegated grey clay	Kerala Ceramics Clay Mine, Kundara
CC-54	Peat	Kerala Ceramics Clay Mine, Kundara
CC-55	Carbonaceous clay	Kerala Ceramics Clay Mine, Kundara
CC-56	Peat	Kerala Ceramics Clay Mine, Kundara

The samples analysed are mainly carbonaceous clays, peat and other clays. The details of samples analysed are given in Table 1. Of the 11 samples macerated, only six proved to be productive (CC28, 51, 53, 54, 55 & 56), while the rest, particularly white and pink clays, were found to be non-polleniferous. Palynomorphs recovered from the samples were processed using the conventional separation methods followed by the ultrasonic cleaning to get rid of the finer particles adhering to the palynofossils. The slides of figured specimens (Figures 3 *a–n*, 4 *a–t*) are deposited in the repository of palynological collections in the Geology and Palaeontology Laboratory for future reference.

The palynoflora recovered from the Bharathi and Kundara clay mines consists of pteridophytic spores, fungal remains and angiospermous pollen grains. Besides, the assemblage contains large number of cuticles, organic spherules and salt glands. The palynotaxa identified are given below:

Fungal remains: Spores and fruiting bodies

Dyadosporonites sp.; *Lirasporis* sp.; *Parmathyrites indicus* Jain & Gupta 1970; *Phragmothyrites eocaenica* Edward 1922.

Pteridophytic spores

Cheilanthoidspora monoleta Sah & Kar 1974; *Cingulatisporis* sp.; *Dandotiaspora plicata* (Sah & Kar) Sah, Kar & Singh 1971; *Osmundacidites cephalus* Saxena 1978; *Polypodiaceaesporites* sp.; *Polypodiisporites* sp.; *Lygodiumsporites* sp.

Angiospermous pollen grains

Compositoipollenites minimus Ramanujam 1987; *Crotonoidaepollenites euphorbioides* Rao & Ramanujam 1982; *Ctenolophonidites costatus* (V. H. Klinkenberg) V. H. Klink. 1966; *Dermatobrevicolporites dermatus* (Sah & Kar) Kar; *Droseridites spinosa* (Cookson) Potonie 1960; *Jandufouria seqmrogiformis* Germeraad,

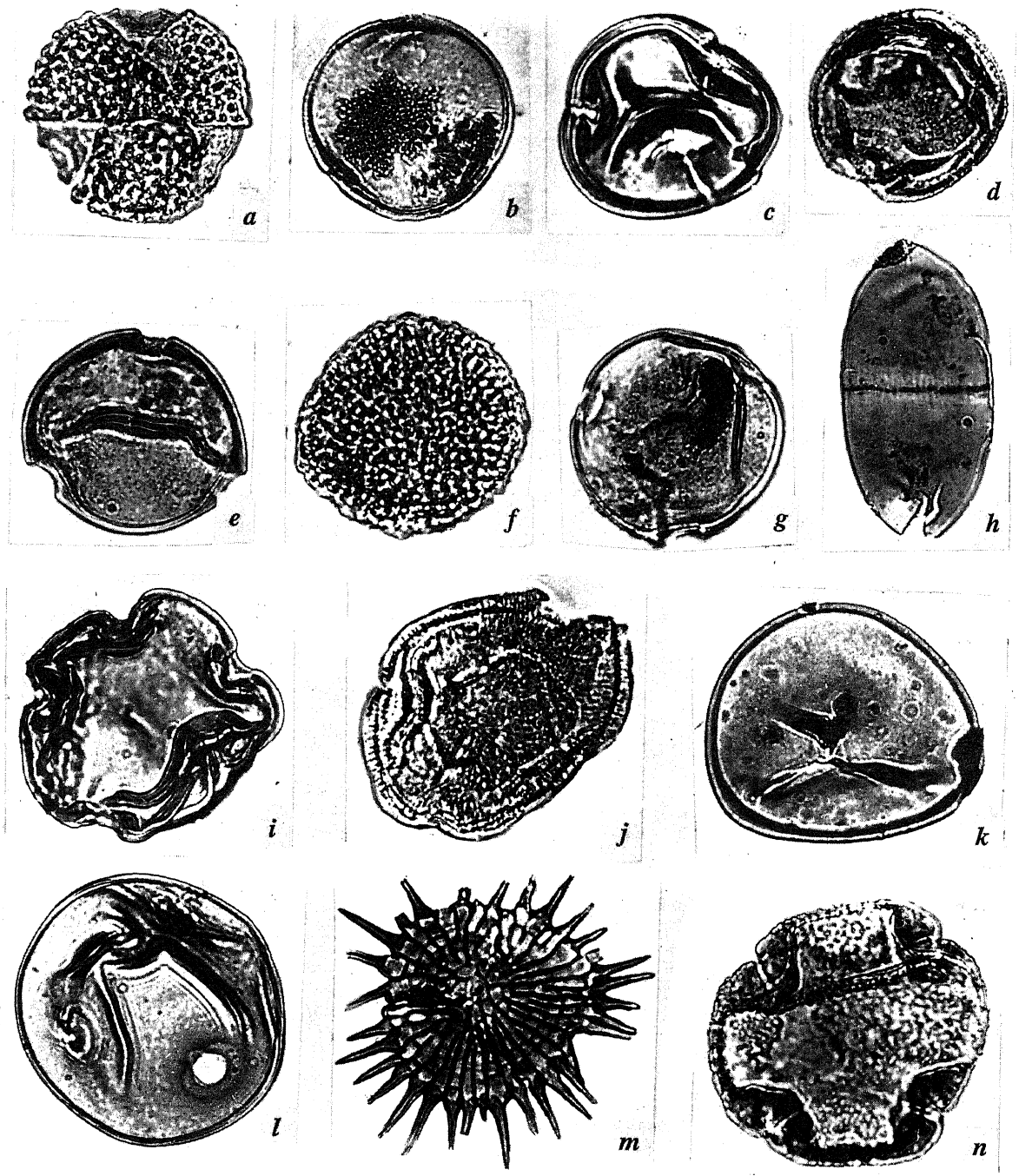


Figure 3. *a*, *Margocolporites tsukadai* Ramanujam. *b*, *Tricolporopollis* sp., *c*, Zonocolpate pollen grain. *d*, *Retitricolporites crassirotatus* Rao & Ramanujam; *e*, *Dermatobrevicolporites dermatus* (Sah & Kar) Kar; *f*, *Crotonoidaepollenites euphorbioides* Rao & Ramanujam; *g*, *Tricolporopollis* Rao & Ramanujam; *h*, *Dyadosporonites* sp.; *i*, *Meliapollis* sp.; *j*, *Jandufouria seamrogoformis* Germeraad, Hopping & Muller; *k*, *Dandotiaspora plicata* (Sah & Kar) Sah, Kar & Singh; *l*, *Lakiapollis ovatus*, Venkatachala & Kar; *m*, *Parmathyrites indicus* Jain & Gupta; *n*, *Meliapollis* sp. cf. *M. tamilii* Navale & Misra. (All photomicrographs are ca × 650).

Hopping & Muller 1968; *Lakiapollis ovatus* Venkatachala & Kar 1969; *Longapertites* sp.; *Malvacearumpollis estelae* (Germeraad, Hopping & Muller) Hekel 1972; *Margocolporites sitholeyi* Ramanujam 1966; *Margocolporites tsukadai* Ramanujam 1966; *Meliapollis tamilii* Navale & Misra 1979; *Meliapollis quadrangularis* (Ramanujam) Sah & Kar 1970; Rhizophoraceae (*Zonocostites*) type pollen grains; *Quilonipollenites* sp. cf. *Q. sahnii* Rao & Ramanujam 1978; *Retitricolporites crassirotatus* Rao & Ramanujam 1982; *Tricolporopollis* sp.; *Trilatiporites* sp.; *Tricolpites crassireticulatus* Dutta & Sah 1970; *Tricolporopilites pseudoreticulatus* Kar 1985; Zonocolpate type.

Of the above-mentioned species, *Margocolporites* are abundant in the assemblage. *Meliapollis* is also fairly represented. Pollen grains belonging to Rhizophoraceae

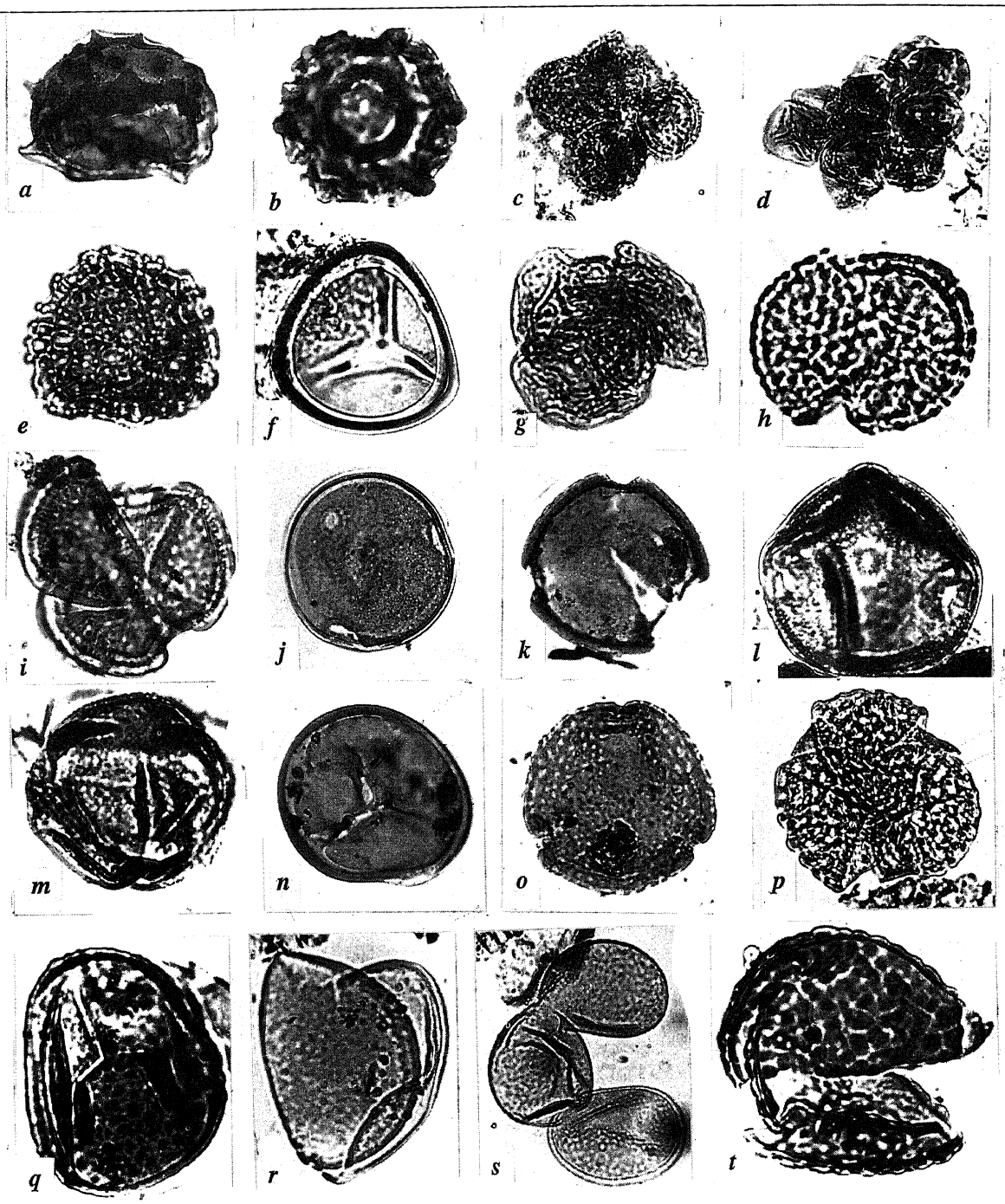


Figure 4. ***a***, *Malvacearumpollis estelae* (Germeraad, Hopping & Muller) Hekel; ***b***, *Ctenolophonidites costatus* sp. (Van Hoeken-Klinkenberg) Van Hoeken-Klinkenberg; ***c***, *Droseridites spinosa* (Cookson) Potonie; ***d***, *Palaeosantalaceaepites* sp. (Group of pollen grains); ***e***, *Cheilanthoidspora monoleta* Sah & Kar; ***f***, *Cingulatisporites* sp.; ***g***, ? Group of Monocolpite finely reticulate pollen grains; ***h***, *Quilonipollenites* sp.; ***i***, *Tricolpites crassireticulatus* Dutta & Sah; ***j***, *Tricolporopollis* sp.; ***k***, *Meliapollis quadrangularis* (Ramanujam) Sah & Kar; ***l***, *Trilatiporites* sp. cf. *T. normei* Ramanujam; ***m***, *Retitricolporites crassirotatus* Rao & Ramanujam; ***n***, *Lygodiumsporites* sp.; ***o***, *Tricolporopilites pseudoreticulatus* Kar; ***p***, *Margocolporites sitholei* Ramanujam; ***q***, *Osmundacidites cephalus* Saxena; ***r***, *Longapertites* sp.; ***s***, *Polypodiaceaesporites* sp. (3 spores together); ***t***, *Polypodiisporites* sp. (All photomicrographs are ca × 550).

(*Zonocostites*) occur mainly in groups attached to cuticles. *Lakiapollis ovatus* is another significant taxon of the assemblage. *Polypodiisporites* and *Polypodiaceaesporites* of Polypodiceae are the important elements of the Pteridophytes. Only some of the forms have been assigned to earlier published taxa, whereas others are referred to informal species.

The occurrence of *Ctenolophonidites, Droseridites, Jandufouria, Lakiapollis, Malvacearumpollis, Margocoloporites, Meliapollis* and *Zonocostites* type of pollen

grain in the assemblage is significant as they are of stratigraphical and palaeocological importance[4–8]. The assemblage as a whole represents a mixture of ecological groups such as low-land, freshwater, mangrove and sandy beach elements, particularly that of the Bharathi clay mine. However, in the Kundara clay mine, the palynofloral spectrum could be differentiated at various levels of sampling. For example, sample CC-56 yielded mainly pteridophytic spores while CC-55 is rich in *Tricolpites crassirotatus*. The occurrence of angiospermous pollen grains namely *Margocolporites tsukadai* and *Meliapollis* spp. has been found to be increasing towards higher levels of sampling in the section. Fungal fruiting bodies are abundant, indicating heavy rainfall and high humidity. The recognition of salt glands in the sediments signifies that the palynoflora has an affinity with mangrove swamp/coastal vegetation. Presence of *Compositoipollenites, Malvacearumpollis, Margocolporites, Meliapollis* spp., *Quilonipollenites* and *Tricolporopilites* indicates that the carbonaceous clays and peat samples belong to the Upper Tertiary[4–11].

The palynofloral assemblage does compare well with the other known contemporaneous deposits of the West coast as they have common floral elements. The palynoflora obtained from the 'Ratnagiri beds' resting on Deccan Trap having a few metre thickness, at Kotwade, Bankhot-Velas and Parchuri of Ratnagiri district of Maharashtra State have several comparable taxa. As this flora[12] was recovered from the deposits which occur at higher altitude of 100–120 m ASL, the assemblage described here comes from a considerably lower altitude, indicating that the former has been subjected to uplift during the post-Miocene age as suggested by several workers[13,14] while the deposits studied here were accumulated in a basin which remained relatively stable as reported earlier[15–17].

Although palynological work has been done covering the Warkalli, the Quilon and the Vaikom formations, a finer resolution of the age of the entire sedimentary sequence of the Kerala basin is yet to be achieved[7]. Most of the published reports on palynoflora are sporadic covering either of one or two formations. Moreover, palynological data of the subsurface sediments are meagre to have meaningful biostratigraphy and to correlate the entire Tertiary sequence overlying the Archean basement. Therefore, it would be worthwhile to analyse more subsurface samples collected from widely drilled boreholes from Kanyakumari in the South to Kasargod in the North for finer resolution of age and a better understanding of palaeoecology and vegetational history of the Kerala coast using palynology as a tool.

Considering the pronounced neotectonic activity along the Maharashtra coast compared to the Kerala deposits, there would be further scope in analysing the palynological evidence for reasonable estimates of dispersal probabilities of several taxa that had been growing along the West coast of India. The palynological evidence *vis-à-vis* present day distribution[18] of *Ctenolophon, Durio, Pentac, Sclerosperma, Syagrus, Nyssa, Elaeis, Gunnera, Kielmeyera, Nypa* and *Eugeissona*, to mention a few, does indicate that ecological stress brought about by tectonic activity might have played a decisive role in their displacement from the West coast of India. However, we have yet to establish the probable reasons for their disjunct distribution and restricted occurrences at oceanic distances as we do not have a comprehensive analysis of pollen data from the West coast. As not much thought is given to the phenological response/readjustment of the species[19] in relation to plate tectonic activity, it would be worth examining the palynological data that are being accrued in that perspective. We have to ponder and adopt a holistic approach[20] so that the role of tectonic activity and the accompanying change in physiography and climatic vicissitudes in geosphere and phytosphere could be understood in a much better way.

1. Soman, K., Machado, T., Joseph, A. and Sreelatha, R. K., Procedings of the Fifth Kerala Science Congress, January 1993, Kottayam, pp. 37–40.
2. Poulose, K. V. and Narayanaswami, S., *Mem. Geol. Soc. India*, 1968, **2**, 300–308.
3. Soman, K., Terry Machado and Ancy Joseph, *Cont. Earth. Sci. Stud.*, Thiruvananthapuram, Tech. Rep., 1992, No. 58, pp. 31.
4. Ramanujam, C. G. K., *J. Palaeontol. Soc. India*, 1987, **32**, 26–46.
5. Ramanujam, C. G. K. and Rao, K. P., *Geophytology*, 1977, **7**, 160–164.
6. Rao, M. R., *Palaeobotanist*, 1990, **38**, 243–255.
7. Singh, H. P. and Rao, M. R., *Palaeobotanist*, 1990, **38**, 256–262.
8. Thanikamoni, G., Caratini, C., Venkatachala, B. S., Ramanujanam, C. G. K. and Kar, K. K., *Inst. Fr. Trav. Sec. Pondicherry Sci. Tech.*, 1984, **19**, 1–92.
9. Kar, R. K. and Bhattacharya, M., *Palaeobotanist*, 1990, **39**, 250–263.
10. Raha, P. K., Rajendran, C. P. and Kar, R. K., *Geophytology*, 1987, **17**, 209–218.
11. Varma, Y. N. R., Ramanujam, C. G. K. and Patil, R. S., *J. Palynol.*, 1986, **22**, 39–53.
12. Badve, R. M., Rajshekhar, C. and Kumaran, K. P. N., Progress Report, Fifth Group Monitoring Workshop on SERC funded projects in Earth Sciences, January 1994, Cochin, pp. 64–71.
13. Rajaguru, S. N. and Marathe, A. R., Proceedings of the Symposium on Quaternary Episodes, Department of Geology, M. S. U., Baroda, 1984, pp. 1–6.
14. Powar, K. B., *Curr. Sci.*, 1993, **64**, 793–796.
15. Agarwal, D. P. and Guzder, S. J., *Palaeobotanist*, 1972, **21**, 216–222.
16. Vaidhyanadhan, R., *J. Geol. Soc. India*, 1987, **29**, 373–378.
17. Caratini, C., Delibrias, G. and Rajagopalan, G., *Palaeobotanist*, 1990, **38**, 370–378.
18. Venkatachala, B. S., *Palaeobotanist*, 1992, **41**, 1–8.
19. Meher-Homji, V. M., *Palaeobotanist*, 1994, **42**, 225–240.
20. Venkatachala, B. S., *Curr. Sci.*, 1990, **59**, 903–907.

ACKNOWLEDGEMENTS. We thank the Directors of both Agharkar Research Institute, Pune, and Centre for Earth Science Studies, Trivandrum, and Dr R. M. Badve for encouragement.

Received 28 August 1995; revised accepted 29 November 1995

Bioremediation of toxic metal ions from polluted lake waters and industrial effluents by fungal biosorbent

Md. Naseem Akthar and P. Maruthi Mohan

Department of Biochemistry, Osmania University, Hyderabad 500 007, India

Alkali-extracted mycelial biomass (biosorbent) from *Aspergillus niger* was effective in sequestering metal ions, especially Zn^{2+} and Cd^{2+} from lake waters from very low concentrations (μg/l) when compared to Dowex-50. Zinc ions (9 g/l) from an effluent of a battery manufacturing unit could be removed efficiently (>75%) by recycling the biosorbent through five cycles. Similarly, chromium ions from effluents of electroplating industry (420 mg/l) could also be removed by recycling the biomass. Here we report the potential use of this fungal biosorbent in removing toxic metal ions from heterogeneously polluted lake waters and industrial effluents.

DEVELOPING countries are increasingly concerned with pollution due to toxic heavy metals in the environment. Unlike most organic pollutants which can be destroyed, toxic metal ions released into the environment often persist indefinitely, circulating and eventually accumulating throughout the food chain, thus posing a serious threat to mankind. Large proportions of heavy metals are being released into the environment from acid mine drains and industrial waste waters due to inadequacies of technology in processing of metals or through other routes[1,2]. A reasonable procedure to diminish the escape of metals from industrial effluents is the adoption of low waste-generating technologies coupled with effective effluent treatment.

The use of biological materials for heavy metal removal or recovery has gained importance in recent years due to their good performance and low cost[3–7]. Among the various sources, both live and inactivated biomass of microorganisms exhibit interesting metal-binding capacities[8,9]. Their complex cell walls contain high content of functional groups like amino, amide, hydroxyl, carboxyl, sulphydryl and phosphate which have been implicated in metal binding[10–14].

Conventional methods for removing metals from industrial effluents before they are disposed off are becoming inadequate, while effective methods are becoming uneconomical[15]. An alternative method, based on the adsorptive capacity of microbial biomass, referred to as biosorption, is emerging as a promising 'biosorption technology' for metal removal and recovery[3,4,16]. Studies from our laboratory had earlier used a nickel-resistant, hyperaccumulating mutant of *Neurospora crassa* to remove toxic Ni^{2+} from aqueous media[17]. In later studies, alkali-extracted mycelial biomass of *N. crassa, Fusarium oxysporum* and particularly *A. niger* have been demonstrated to have an excellent ability to concentrate silver ions[10]. Further, the mechanism of biosorption was demonstrated to be by a direct exchange of toxic ions with resident Ca^{2+}/Mg^{2+} of the biosorbent; carboxyl groups have been shown to be the major binding groups herein[11]. The objective of the present study is to explore the potential use of a biosorbent from *A. niger* in removing toxic metal ions from polluted lake waters and industrial effluents.

The biomass of *A. niger* was processed as described earlier[10], referred to as 'biosorbent' was used in all experiments. For metal-removal experiments, Hussain Sagar, an eutropically polluted lake in the Hyderabad city, India, was chosen and the metal ion concentrations (Cd^{2+}, Cu^{2+}, Co^{2+}, Ni^{2+} and Zn^{2+}) were determined[18]. The biosorbent (25 g fresh weight = 5 g dry weight) was taken in a cheese cloth bag and dipped into the lake waters (2–3 feet deep) for 60 min. The biosorbent was then removed, washed with distilled water and the bound metals were eluted with 0.1 N HCl. The eluate was concentrated and the amount of metals bound was anayysed by Atomic Absorption Spectrophotometry (AAS, Perkin–Elmer Model No. 2380). Simultaneously for comparison of efficiency of the biosorbent, 5 g of Dowex-50 (Na^+ form) was taken in a cheese cloth bag and dipped into the lake waters and the experiment was carried out as described above for biosorbent.

For the removal of metals from industrial effluents, biosorbent (2 g) was suspended in 20 ml of the effluent and incubated on a rotary shaker (100 rpm) at 30 ± 1°C for 1 h and the amount of metal remaining was analysed. The biosorbent was reused for 5 cycles after regenerating with Ca^{2+} and Mg^{2+} (0.1 M each) as described earlier[10].

Microorganisms have applications in selected wastewater metal removal/treatment systems. However, toxicity of metallic contaminants and the highly variable conditions prevalent in many wastes and processed waters preclude the use of living organisms and necessitate the utilization of non-living systems for metal removal. Metal sorption has been well demonstrated with nonmetabolizing and nonliving, microbial biomass[16,19,20]. Several characteristics of inactive biomass make it an ideal choice for treatment of metal containing waste waters and process streams[21,22]. Inactive biomass, in general, is not selective in the metal sorbed; rather, it simultaneously removes several different toxic and heavy metals like Cd, Cr, Cu, Hg, Ni, Pb, U, Zn from solution regardless of their concentrations. In addition, the biomass has greater avidity to and removes metals that are relatively more hazardous, due to its low affinity for alkaline earth metals (Ca, Na, K, Mg) (ref. 23).

Most studies described above have been conducted using metal ions in pure solutions and very few have

Table 1. Sequestration of toxic metal ions from lake waters.

Metal ion	M^{2+} in lake water µg/l*	Amount of metal bound µg/5 g		Litre** equivalents
		Biosorbent	Dowex-50	
Cd^{2+}	8.63 ± 6.5	333.8 ± 155.5	8.69 ± 4.97	38.7
Cu^{2+}	18.1 ± 14.5	85.1 ± 41.2	21.94 ± 9.3	4.7
Zn^{2+}	62.3 ± 21	4562.5 ± 1174.3	111.9 ± 67.3	73.2
Ni^{2+}	17.98 ± 7	90.13 ± 52.48	6.35 ± 3.37	5.3
Co^{2+}	10.65 ± 1.1	84.69 ± 36.75	7.75 ± 2.95	7.9

*Five samples were collected from the same site of the lake over a period of one year and pH varied between 7.9 and 8.5.
**Obtained by dividing biosorbent-bound metal ion by the lake water metal concentration per litre.
Biosorbent or Dowex-50 (5 g) placed in a cheese cloth bag and immersed into lake water (2–3 feet deep) for 60 min. The bound metal ions were desorbed and estimated by AAS (see text).

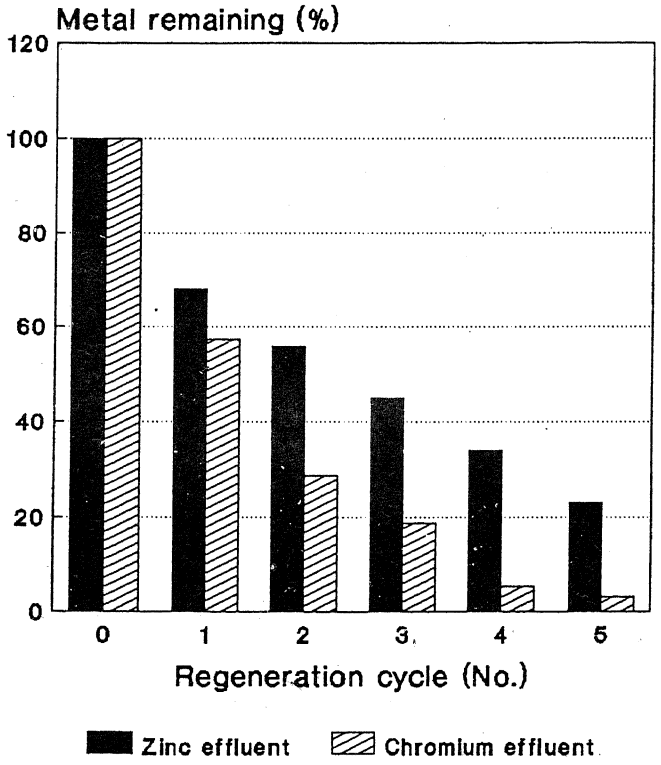


Figure 1. Biosorbent (2 g) was suspended in 20 ml of effluent containing zinc (9 g/l) or chromium (420 mg/l) for 60 min. The metal ion remaining was estimated by AAS. After regeneration, the bisorbent was reused up to 5 cycles (see text).

used industrial effluents to our knowledge[24,25]. In the present study, we have used the polluted effluents directly without manipulating their composition. The results show that cadmium and zinc were more effectively sequestered from the lake waters at very low concentrations (ppb or µg/l) as compared with other metal ions by the biosorbent. The efficacy of the biosorbent *vis-à-vis* equal weight of Dowex-50 indicated the former's superiority in binding metal ions (Table 1). This is due to Dowex-50 readily binding nonspecifically divalent ions present in lake waters (like Ca^{2+} and Mg^{2+}), thereby attaining saturation with these ions (data not shown) and being there after unable to sequester toxic ions present at very low concentrations. If the binding capacity of the biosorbent is calculated on the basis of litre equivalents, the biosorbent (5 g) is able to remove almost 75 l equivalents of zinc and around 40 l equivalents of cadmium: for copper, cobalt and nickel, equivalence was around 5–10 l (Table 1). The metal ion concentrations noted here are well below the permissible limits suggested by the Indian Council of Medical Research[26]. It should be noted here that when the same metal ions were present in a mixture at high concentrations (25 mM each), the order of binding by the biosorbent (from pure solutions) was Cu > Zn > Cd > Co = Ni (ref. 11). This shows that other unknown factors influence the order of preference for metal sorption from such heterogeneously polluted waters.

Zinc removal from an industrial effluent (from a battery manufacturing unit) which contains a high concentration of zinc (9 mg/ml) was highly effective with the biosorbent; five cycles of biosorbent treatment resulted in more than 75% metal removal. In another effluent from an electro plating industry containing chromium (420 µg/ml), the biosorbent was able to remove more than 95% of chromium (five cycles). One of the interesting features regarding Zn^{2+} sequestration from effluent of a battery manufacturing unit is that in the first cycle of use (Figure 1) about 30% of Zn^{2+} is removed by the biosorbent (60 mg zinc/2 g fresh wt). On a dry weight basis, this amounts to 15% of biosorbent weight (as dry weight accounts for 20% of fresh weight). In earlier studies[11], Zn^{2+} binding was 9–10% from a pure solution containing 50 mM (about 3 mg/ml). The higher binding noted herein is possibly due to a 3-fold higher concentration of Zn^{2+} present in the effluent. The accumulation of toxic metal ions is based on a mole for mole exchange with the resident calcium and magnesium ions of the biosorbent. The low affinity of Ca^{2+} and Mg^{2+} ions for the biosorbent makes them excellent counter ions for the heavy metals that form more stable complexes with the unprotonated binding sites. This factor has been taken as an advantage for the regeneration of the biosorbent and its repeated use[11].

Ideal properties for a biosorbent are: i) it should sequester toxic metal ions both from dilute and concentrated solutions, ii) should not face antagonism with essential metal ions, iii) should be reusable, iv) should not by themselves contaminate the environment. Overall, the biosorbent used in the present study has most of these features. Further scale-up studies are required to determine the actual potential of this biosorbent for decontaminating metal ions from industrial effluents.

1. Volesky, B., *FEMS Microbiol. Rev.*, 1994, **14**, 291–302.
2. Fourest, E., Canal, C. and Jean-Claude, R., *FEMS Microbiol. Rev.*, 1994, **14**, 325–332.

3. Volesky, B., *TIBTECH*, 1987, **5**, 96–101.
4. Gadd, G. M., in *Microbial Control of Pollution*, (eds Fry, J. C., Gadd, G. M., Jones, C. W., Slabert, R. A. and Watson–Craik, I.), Cambridge Univ. Press, Cambridge, 1992, pp. 59.
5. Brierley, C. L., *Geomicrobiol. J.*, 1990, **8**, 201–223.
6. Macaskie, L. E., *J. Chem. Tech. Biotechnol.*, 1990, **49**, 357–379.
7. Mattuschka, B. and Straube, G., *J. Chem. Tech. Biotechnol.*, 1993, **58**, 57–63.
8. Kuyucak, N. and Volesky, B., *Biotechnol. Lett.*, 1988, **10**, 137–142.
9. Siegel, S. M., Galun, M. and Siegel, B. Z., *Water, Air, Soil Pollut.*, 1990, **53**, 335–344.
10. Akthar, N., Sastry, K. S. and Mohan, P. M., *Biotechnol. Lett.*, 1995, **17**, 551–556.
11. Akthar, N., Sastry, K. S. and Mohan, P. M., *Bio Metals*, 1995, in press.
12. Birch, L. and Bachofen, R., *Experientia*, 1990, **46**, 827–833.
13. Treen-Sears, M., Volesky, B. and Neufeld, R. J., *Biotechnol. Bioeng.*, 1984, **26**, 1323–1329.
14. Kuyucak, N. and Volesky, B., *Biotechnol. Bioeng.*, 1988, **33**, 809–814.
15. Brauckmann, B. M., in *Biosorption of heavy metals* (ed. Volesky, B.), 1990, CRC Press, Boca Raton, FL, pp. 51–64.
16. Paknikar, K. M., Palnitkar, U. S. and Puranik, P. R., in *Biohydrometallurgical technologies* (eds Torma, A. E., Apel, M. L. and Brierley, C. L.), vol. II, pp. 229–235.
17. Kumar, C. S., Sastry, K. S. and Mohan, P. M., *Biotechnol. Lett.*, 1992, **14**, 1099–1102.
18. Radha, S. and Seenayya, G., *The Sci. of the total Environ.*, 1992, **125**, 123–136.
19. Hutchins, S. R., Davidson, M. S., Brierley, J. A. and Brierley, C. L., *Annu. Rev. Microbiol.*, 1986, **40**, 311–316.
20. Brierley, C. L., in *Microbial Mineral Recovery* (eds Ehrlich, H. L. and Brierley, C. L.), 1990, McGraw-Hill Publishing company, New York, pp. 303–323.
21. Wales, D. S. and Sagar, B. F., *J. Chem. Tech. Biotechnol.*, 1990, **49**, 345–355.
22. Gale, N. L., *Biotechnol. Bioeng. Symp.*, 1986, **16**, 171–180.
23. Wainwright, M., An Introduction to fungal Biotechnology, John Wiley & Sons, Chichester, UK, 1992, 94–101.
24. Lewis, D. A. and Kiff, R. J., *Environ. Tech. Lett.*, **9**, 991–998.
25. Leuf, E., Prey, T. and Kubicek, C. P., *Appl. Microbiol. Biotechnol.*, **34**, 688–692.
26. ICMR, Indian Council of Medical Research, New Delhi, India, 1975, special report, pp. 44.

ACKNOWLEDGEMENTS. We acknowledge the financial assistance from the Department of Science and Technology and COSIST grant from University Grants Commission, New Delhi; the Department of Biochemistry, Osmania University. We thank Prof. K. Sivarama Sastry for reading the manuscript.

Received 4 July 1995; revised accepted 7 November 1995

Deep Continental Structure of India: A review. T. M. Mahadevan, Geological Society of India, P.B. No. 1922, Bangalore 560 019. Memoir 28. 1994. 569 pp. Price: Rs 600/$60.

The book under review is a monumental compilation of geophysical field research on deep continental structure of India. A short preface summarizes the scope and objectives and expresses the hope that this work will help to bridge a communication gap between geologists and geophysicists, for whom deep continental studies are common ground. The text begins with (1) 'General Considerations', which briefly defines the Indian continental crust and its geographical extent along with geological and geophysical characteristics, particularly of its lower crust, Moho and upper mantle. In Chapter 2, 'Technology and Methods – Present Status', a short description of the geophysical techniques for imaging the deep interior is highlighted, with emphasis on standards in technology. The contents of this chapter would make the reader feel that a comparative attempt has not been made of the synthesized data from different methods for the Indian lithosphere or any of its seven tectono-stratigraphic units specified in the next chapter. Chapter 3, 'The Indian Crust', deals with the results of various geophysical investigations of the seven major geological entities. Here the author has summarized the informations of velocity structures, heat flow, Himalayan isostasy, etc., without their critical evaluation. Chapter 4, 'Peninsular Shield – Precambrian Cratons and Mobile Regions', gives some broad features of geology and deep crust. A vital problem is emphasized as to whether the Indian plate has evolved by collision and fusion of a number of microplates or it was a single ancient continental crust that underwent intermittent rifting to form sedimentary basins which finally developed into foldbelts now seen bordering the cratons. In Chapter 5, 'Purana Basins', the author considers these basins as Proterozoic analogues of the Tertiary basins that occur in front of the Himalayan orogen, although the available geophysical information does not support, except for the Cuddapah basin. In 'Gondwana Basins', described in Chapter 6, the horst and graben structures are revealed in gravity and DSS profiles. Despite the shallow basin bottoms, some basins like Godawari and Mahanadi are interpreted as zones of intra-continental passive rifts without establishing their Precambrian ancestry. Chapter 7, 'Mesozoic-Tertiary Volcanic Province', gives an account of all possible evolutionary models but the geophysical inputs have not been always used as evidence in support or against any of these models. The author has followed a simple theory where positive gravity field is attributed to rising Moho or to the presence of basic/alkaline rocks, while negative field is taken to indicate sediments. The Sonata belt is elaborately described but this reviewer feels that no definite conclusions can be drawn from the geophysical data alone for the evolution of this major tectonic divide of Sonata. In Chapter 8, 'Younger Sedimentary Basins', a significant feature is their horst and graben structures. The Cambay basin is classed as a failed arm with Kutch basin and Sonata belt as other two arms of a triple junction, akin to the Afar triangle from which three rift systems radiate. But this hypothesis would be misleading to earth scientists as the geological and geophysical methods have not yet been evaluated in this direction. The section on geomagnetic field and conductive structures on page 327 appears somewhat out of place in this chapter. In Chapter 9, 'Himalaya', the gravity and seismic data have been synthesized with surface geology, but different evolutionary models are simply stated. The reader is, however, left to himself to relate the geophysical data with the vertical tectonics or plate tectonics that have been responsible for the Himalayan orogeny. While not competent on Chapter 10, 'Continental Margins of India', I am confident that this will be welcomed as a reliable compilation. In Chapter 11, 'Emerging Scenarios of Deep Continental Structures – A Resume', an attempt is made to synthesize the various geophysical data on deep continental crust of India. Considering thickness as most useful parameter, the Moho depths and lithospheric thickness are summarized in tables. The review is a worthy attempt and quite elegantly examines models already presented without their critical evaluation. The last Chapter 12. 'Framework for Future Programme', contains useful recommendations, although with a long list. A multidisciplinary, integrated and careful planning is recommended for future researches on deep continental structures of India. There is truly magnificient list of about 900 references.

The vast geophysical information contained in this volume will prove useful in establishing deep crustal interaction with surface geology. The text is clear and on the whole very readable, except when some of the arguments and statements appear to be out of context and when long sentences make severe demands upon the reader's power of concentration. The book is not without minor flaws. Some figures are nearly illegible or fuzzy because of extreme reduction, a few illustrations are incorrectly labelled and the text has a few typographical errors. However, the volume contains so much valuable material assembled in so useful a form that one can hope it will be widely disseminated and read. One must salute the courage and patience of the author who, being a geologist with vast experience, ventured for the first time in reviewing the massive geophysical work on a request from the Department of Science and Technology, New Delhi, which initiated modern research activities on Indian deep continental crust.

RAM S. SHARMA

Department of Geology,
Banaras Hindu University,
Varanasi 221 005, India

Rice Tungro A. Anjaneyulu, M. K. Satapathy and V. D. Shukla. Oxford and IBH Publishing Co. Pvt. Ltd., New Delhi, Bombay, Calcutta, 1994, 228 pp.

Rice is a staple food for human beings. Rice tungro virus is considered as one of the major constraints in rice production in South East Asia. For many years plant virologists, plant pathologists and entomologists have done research on tungro virus disease. Prior to this publication comprehensive review was not available. This book fulfils the needs of the researchers and the students who wish to gain a fair degree of knowledge on tungro virus disease.

The text has been divided into 19 chapters. Chapter one presents an overview of tungro virus disease. The occurrence, distribution and area affected in different parts of the world have been clearly indicated in the map (chapter 2). Symptoms are important in any disease

diagnosis. Range of symptoms by different strains of virus on different rice cultivars has been described. Symptoms induced by feeding vector leafhoppers are also given (chapter 3). The management strategy depends upon yield loss. Though yield reduction on different varieties has been given yield loss per unit area is not provided (chapter 4).

For many years the cause of the disease was not known. Rice tungro is caused by bacilliform particles of 25 x 140 nm and spherical particles of 28–30 nm in diameter. Antisera has been raised against both the particles of tungro virus. This is most useful in the detection and epidemiology. Various diagnostic techniques have been described in chapters 5 and 6.

Molecular biology of tungro virus is most interesting because of two viruses; RNA virus and DNA virus associated in causing the disease. Only limited information is available on this (chapter 7). Existence of several strains of RTV (RTV 1, RTV 2, RTV 2A, RTV 2B, RTV 3, RTV 4) in different rice-growing areas has been described (chapter 8). Little attention has been paid to the metabolic processes associated with the symptom development of rice plants infected with tungro viruses despite its widespread occurrence and economic importance (chapter 9).

Tungro viruses are transmitted by various species of *Nephotettix*. The authors have described vector distribution, virus–vector relationships, biology, host range and vector movement (chapters 10–12). The epidemiology of tungro disease is interesting because of the involvement of insect vector, besides host (rice), pathogen (RTSV & RTBV) and environment. Presence of virus strains, vector biotypes, multiple rice cultivars, alternate weed hosts and different cultural practices make tungro epidemiology very interactive and disease development becomes a function of various aspects of virus, host, agronomic practices and materiological factors (chapter 13). Tungro viruses, their transmission, virus–vector relationship, vector biology, population dynamics of the vector and host resistance have been discussed (chapters 10–13). Various ways and means have been attempted and different recommendations have been made in different Asian countries to reduce the disease incidence and increase the crop yields (chapter 14). Varietal resistance has been discussed in a separate chapter (chapter 15). An interesting aspect is a note for farmers and extension workers provided in the book (chapter 17), besides future prospects (chapter 18). All the literature published on rice tungro is provided in the book (chapter 19) and subject index is given for easy reference.

The book is an outstanding contribution with application of data, diagrams and illustrations. I hope it will be useful to plant pathologists, plant virologists, entomologists, extension workers and students and teachers of general agriculture.

V. MUNIYAPPA

Department of Plant Pathology, University of Agricultural Sciences, GKVK, Bangalore 560 065, India

Bacterial Pathogenesis of Plants and Animals: Molecular and Cellular Mechanisms. J. L. Dangl, ed. Current Topics in Microbiology and Immunology series, Springer-Verlag, Berlin, Heidelberg. 1994. Vol. 192. 343 pp.

Although four highly authoritative books (*Molecular Basis of Bacterial Pathogenesis* eds Barbara H. Iglewski and Virginia L. Clark); *Molecular Biology of Bacterial Infection: Current Status and Future Perspectives* (eds Hormaeche, C. E., Penn, C. W. and Smyth, C. J.); *Bacterial Pathogenesis: A Molecular Approach* (eds Abigail A. Salyers and Dixie D. Whitt); *Molecular Genetics of Bacterial Pathogenesis* (eds Miller, V. L., Kaper, J. B., Portnoy, D. A. and Isberg, R. R.). For reviews see *Curr. Sci.*, 1993, **64**, 118–119; 1993, **65**, 798–799; 1995, **68**, 343–344; 1995, **69**, 624–625) have already appeared on the molecular underpinnings of bacterial pathogenesis, the slant of this volume is entirely different. This is apparent, at the outset, from the somewhat non-obvious choice of topics which include both plant and animal pathogens. The objective of this volume is to bring out the commonalty of molecular mechanisms which underlie bacterial pathogenicity, irrespective of the type of the pathogens – plant or animal, and thus underscore the importance of basic research approaches to parallel problems in pathogenesis.

The first basic concept about bacterial pathogenesis, whether plant or animal, which one realizes while going through this volume is the multiplicity of virulence factors required. Beattie and Lindow discuss various virulence factors which determine the epiphytic fitness of phytopathogenic bacteria. This is analogous to attachment and colonization of host surfaces by the animal pathogens as detailed by Meyer *et al.* and Manning in their reviews on the pathogenicity of *Neisseria* spp. and *Vibrio cholerae* respectively. From these reviews, it is evident that while not much is known as to the molecular mechanisms underlying various virulence factors, viz. fimbrial adhesions, extracellular polysaccharides (EPS), osmotolerance, tolerance to UV and visible radiations, etc. determining the epiphytic fitness of phytopathogenic bacteria, detailed information is available on the molecular genetics of colonization factors of animal pathogens such as Tcp (*t*oxin *c*oregulated *p*ilus) and Acf (*a*ccessory *c*olonization *f*actor) of *Vibrio cholerae*, and pili and opa (*op*acity *a*ssociated) proteins of *Neisseria* spp.

The next basic concept which one realizes is the complexity of genetic regulatory pathways controlling expression of virulence determinants in bacterial pathogens. Dow and Daniels discuss global regulation of virulence determinants, viz. extracellular enzymes (cellulase, protease, polygalacuronase lyase, amylase and lipase), EPS, xanthan, etc. of *Xanthomonas campestris* pv. *campestris*. This is a clear reflection, among animal pathogens, of *bvg* locus of *Bordetella pertussis* as detailed by Rappouli. The review by Dow and Daniels also shows that environmental modulation of virulence determinants in phytopathogenic bacteria, as is well known for the animal pathogens, is also regulated by two-component signal transduction mechanisms. The existence of two-component regulatory systems in phytopathogenic bacteria is further corroborated by studies on the genetic and chemical bases of recognition in *Agrobacterium tumefaciens* as reviewed by Binus and Howitz.

The extracellular products produced by both animal and plant bacterial pathogens represent a group of 'most obvious' virulence determinants. Collmer and Bauer, choosing two quite different plant pathogens namely *Erwinia chrysanthemi* and *Pseudomonas syringae*, discuss the molecular genetics of extracellular virulence proteins. In doing so the authors put maceration, pectic enzymes and *pel* genes *vis-a-vis* hyper-

sensitive response (HR), hairpins and *hrp* genes. The concluding section of this review is very interesting in which the authors compare pathogenicity of *Erwinia chrysanthemi* with that of *Pseudomonas aeruginosa* and *P. syringae* with *Yersinia pestis* (reviewed in this volume by Cornelis), thus bringing out similarities/dissimilarities in the deployment and functions of their extracellular virulence proteins. The review by Cornelis is a detailed account of the molecular genetics of yops, the extracellular virulence proteins produced by *Yersinia pestis*.

An extensive review by Bonas discusses structural organization and functions of *hrp* (*h*ypersensitive *r*eaction and *p*athogenicity) genes and *hrp*-dependent secretion of hypersensitive response-inducing proteins (hairpins) in *Erwinia amylovora* and *Pseudomonas syringae* pv. *syringae*. Among the animal pathogens this system can be compared, in the larger perspective, to a number of virulence genes of *salmonella* (reviewed by Finlay). The various genes of *Salmonella* which are under intensive investigation are: *Spv* (*S*almonella *p*lasmid *v*irulence) *Spa* (*S*almonella *p*resentation of *a*ntigen) and *pag* (*Pho-P a*ctivating *g*enes).

The commonalty of the molecular and cellular mechanisms among bacterial pathogens producing quite different clinical disease entities is illustrated by the study of *Shigella flexneri* and *Listeria monocytogenes* as reviewed by Parrot and Sheehan *et al.* respectively. Although quite different as to the clinical disease produced, these pathogens show several similarities in their cytoplasmic multiplication, actin-based movements and cell-to-cell spread.

In a detailed review entitled 'enigmatic avirulence genes of phytopathogenic bacteria' Dangl discusses whatever little is known about the organization, regulation and virulence functions of avirulence (*avr*) genes, bringing into focus some of the paradoxical issues of the topic and posing several questions which must be explored further to understand their exact role in pathogenesis.

A very important question which emerges from the foregoing discussion is: taking cues from animal pathogens, where molecular genetics of pathogenicity are decidedly better understood, can we address ourselves to some of the hitherto unsolved issues of plant pathogens or even vice-versa? This question has been posed more convincingly by McKhann and Hirch in their review on *Rhizobium*. Arguing from several points of view the authors show that there are important conceptual differences between symbiotic and plant pathogen interactions and suggest that *Rhizobium*-legume symbiosis cannot be regarded as a modified or refined type of plant pathogen interaction as proposed by some investigators sometimes back. They further suggest that to understand root hair invasion by rhizobia we should, perhaps, look into models representing invasion of mammalian cells by bacterial pathogens.

The objective of this intellectually stimulating volume is to make the readers view bacterial pathogenicity from newer perspectives. Seen in this context, the volume in its totality is a laudable attempt and a novel addition to works already available on the subject. However some of the reviews, when seen individually, tend to look aloof and wanting in overall objectives of the compilation. This apart each review, whether dealing with plant or animal pathogens, gives detailed information on the molecular and cellular mechanisms underlying the bacterial pathogen discussed. Each review has been provided with an extensive and up-to-date bibliography.

This volume is strongly recommended as part of any collection of books dealing with microbiology. The volume will be of much use to researchers studying various aspects of the molecular genetics of bacterial pathogens and to teachers engaged in teaching advanced level courses on bacteriology plant or animal bacterial pathogenesis and infectious diseases.

J. S. VIRDI

Department of Microbiology,
University of Delhi South Campus,
Benito Juarez Road,
New Delhi 110 021, India

Erratum

Pioneers of ozone study bag chemistry Nobel

S. Parthiban

[*Curr. Sci*,. 1995, **69**, 793–797]

The photographs of Rowland and Molina were inadvertently interchanged. The error is regretted.

– Editors

Edited and published by Prof. S Ramasheshan and Prof. P. Balaram, Current Science Association, Bangalore 560 080.
Typeset by WINTECS, Bangalore 560 021, Printed at Printek Printers, Bangalore (3357763).

INSTITUTE FOR PLASMA RESEARCH

(An Autonomous Institute of DST, Government of India)
Near Indira Bridge, Gandhinagar Highway,
BHAT, GANDHINAGAR 382 424

ADVT NO. 09/95

FACULTY AND POST DOCTORAL POSITIONS IN SURFACE AND MATERIAL SCIENCES

The Institute for Plasma Research is the country's premier research institute pursuing research and development in plasma physics and its applications. The institute has been established by the Department of Science and Technology, Government of India to lead the country's Plasma Physics and Magnetic Confinement Fusion Research Programme.

Tokamak 'Aditya' is in operation and a steady state superconducting tokamak (SST) is under design. Other areas of thrust are basic plasma physics, fusion technologies, plasma-assisted material processing, computer simulation, etc. Improved understanding of plasma–surface interaction is important from the point of view of design and engineering of magnetic confinement fusion devices and plasma-assisted material processing. In the former the degradation of integrity of plasma facing components and the control of evolution of impurities are the issues of critical importance, while in the latter, technologies like surface engineering are based on our understanding of plasma–surface interaction.

IPR is planning to strengthen its activities in this field. A full-fledged surface and materials characterization laboratory with advanced instrumentation like ESCA, XPS, AES, SIMS, SEM, XRD, etc. is being established.

We are looking for scientists at the faculty level like Associate Professor – I & II, Fellow, Scientist – SD and Post-Doctoral Fellow to initiate and lead the programme at IPR. While the thrust will be on goal-oriented research and development mentioned above, individual scientists will have considerable freedom to pursue their own special interests.

FACULTY AND POST DOCTORAL POSITIONS

Qualification and Experience: Ph D with adequate research experience in material and surface analysis and characterization using the above-mentioned techniques.

Job description: Research and development in surface and material science, surface and material analysis and characterization using advanced analytical techniques.

GENERAL INFORMATION

Scale of pay:

1. Associate Professor – II: Rs 4500–150–5700 (gross Rs 10,910)
2. Associate Professor – I: Rs 3700–125–4700–150–5000 (gross Rs 9830)
3. Fellow/Scientists – SD: Rs 3000–100–3500–125-4500 (gross Rs 8180)
4. Post-Doctoral Fellow (PDF): Consolidated stipend Rs 4000 per month + Rs 1000 House Rent Subsidy.

Higher starting salary in the scale may be considered in exceptional cases.

All posts are regular except that of PDF which is for one year tenure which may be extended for another term

Age limit: 35 Years in respect of Fellow/Scientist-SD and PDF. Relaxable by 5 years in case of SC/ST candidates.

The above norms may be relaxed in exceptional cases at the discretion of competent authority. Institute offers excellent career opportunities to creative, innovative and result-oriented staff. In addition to pay and allowances, perquisites like Contributory Provident Fund, Gratuity, Medical Facility under its Contributory Health Service Scheme, LTC, Group Linked Insurance Scheme, Subsidised Transport/Conveyance Reimbursement and Subsidized Canteen Facilities are admissible as per Institute rules.

Out-station candidates called for interview will be paid first class to and fro rail/ordinary class bus fare by shortest route in respect of Associate Professor – II & I, Fellow/Scientist-SD and second class fare for PDF on production of proof of journey such as Railway/Bus ticket or ticket No.

How to apply: Typed application on plain paper strictly in the following format with one recent passport size photograph duly affixed on it along with copies of marksheets, experience certificates, testimonials, etc. should reach the Chief Administrative Officer, Institute for Plasma Research at the above address within 20 (twenty) days from the date of publication of this advertisement. **The envelope should be strictly superscribed with Advt No. and Post Applied for.**

Application format

1. Advertisement No.; 2. Post applied for; 3. Name (in block letters); 4. Date of birth; 5. Place of Birth; 6. Age; 7. Marital status; 8. Present address; 9. Permanent address; 10. Whether belongs to SC/ST if so, attach documentary evidence; 11. Qualifications (SSC onwards, indicating a. exam. passed, b. board/university, c. subjects, d. year of passing, e. division, f. percentage of marks obtained); 12. Experience (indicating a. name of the organization, b. nature of work in detail, c. period, d. scale of pay and total salary drawn); 13. References 14. Any other relevant information; 15. Signature with date.

Candidates employed in Govt/quasi Govt/Public Sector undertakings should forward their applications through proper channel. The Institute reserves the right to reject any application without assigning any reason. **Applications not in the above format are liable to be rejected.** No interim correspondence will be entertained.

Current Science

Special Publications

Title	Guest Editor	Price India (Rs)	Price Abroad (US$)
1. *S. N. De and Cholera Enterotoxin*	P. Balaram	50	12
2. *The Ramachandran Map and Protein and Peptide Structures*	P. Balaram	50	12
3. *Waves and Symmetry*	G. Srinivasan	50	15
4. *A Celebration of Colour in Astronomy*	G. Srinivasan	175	20
5. *Clinical Research and Health Care in Developing Countries*	G. N. Menon	50	12
6. *Biotechnology in India* (copies out of stock)	P. Balaram	–	–
7. *Remote Sensing for National Development* (copies out of stock)	M. G. Chandrasekhar, B. L. Deekshatulu, V. Jayaraman, George Joseph, R. L. Karale, K. Kasturi-rangan, S. Krishnamurthy, K. Radhakrishnan	–	–
8. *Extinct Plants, Evolution and Earth's History*	B. S. Venkatachala, C. V. Subramanian	70	20
9. *Seismology in India – An Overview*	Harsh K. Gupta	100	20
10. *Alzheimer's Disease: An Emerging Issue for the Developing Countries*	George M. Martin	150	15
11. *Artificial Intelligence*	R. Narasimhan	100	12
12. *Quaternary Period in India*	K. S. Valdiya	150	20
13. *Pollination Biology in Tropics*	K. S. Bawa, K. N. Ganeshaiah, R. Uma Shaanker	100	12
14. *Special Issue on Inner Planets*	R. N. Singh	150	25
15. *Condensed Matter Science – Current Status and Plan for Action*	T. V. Ramakrishnan S. V. Subramanyam	150	20
16. *Neuroscience – A multidisciplinary approach*	P. N. Tandon Raghavendra Gadagkar	150	20
17. *Reproductive Biology*	T. C. Anand Kumar	150	20

Contact: The Circulation Department, Indian Academy of Sciences, P. B. No. 8005, Bangalore 560 080, India.

Tel.: 91-80-3342546; Grams: ACADEMY; Telex: 845-2178 ACAD IN; Fax: 91-80-3346094.

Index to Current Science, Volume 69, July–December 1995

AUTHOR INDEX

Abdussattar 924 (Research communication)
Abraham, Oomen 396 (Correspondence)
Absar, N. *see* Sarker, A. A.
Achaya, K. T. 415 (General article)
Adak, T. 695 (Research communication)
Agarwala, B. K. 587 (Review article)
Ahmad, S. Masood 691 (Research communication)
Ahuja, P. S. *see* Giri, C. C.
Ajitha, T. M. *see* Rao, P. V. S.
Ajmani, Ranjeet S. 877 (Research communication)
Akthar, Md. Nam *see* 1028 (Research communication)
Aman, Ashutosh 772 (Research communication)
Ananth, P. N. *see* Tatavarti, Rao
Ansari, Habeebul Haque 367 (Book review)
Ansari, M. A. *see* Adak, T.
Arumugam, M. 21 (Scientific correspondence)
Arunachalam, Subbiah 285, 485 (News); 548 (Book review)
Arvind Kumar 702 (Book review)
Arya, S. K. 295 (Scientific correspondence)
Athavale, R. N. 279 (Historical commentary and notes)
Atul Kumar *see* Kumar, Vandana A.
Au, Gary K. 499 (General article)
Awasthi, D. D. 371 (Book review)

Badarinath, K. V. S. 604 (Research communication)
Bagchi, Biman 129 (Review article)
Bahr, D. *see* Satyam, P. V.
Bains, N. S. 621 (Research communication)
Bal, Madhusmita 538 (Research communication)
Balakrishna, P. 926 (Research communication)
Balaram, P. 7, 301 (Editorial notes); 84, 211, 284, 391, 631, 715 (In this issue)
Balaram, V. 640 (General article)
Balasubramanian, A. S. *see* Satish, P. R.
Banerjee, Rahul *see* Maithani, P. B.
Banerjee, Rakhee *see* Singh, B. N.
Banerji, Subrata 252 (Research article)
Banubakode, P. D. *see* Srivastava, A. K.
Basu, A. N. *see* Sen, D.
Bavle, Amar 491 (Research news)
Bayer, Ronald 861 (Article, special section)
Bedi, J. S. 21 (Scientific correspondence)
Beuria, M. K. 257 (Research communication)
Bhat, Sushma 607 (Research communication)
Bhatia, C. R. 963 (News)
Bhatia, Rajendra 69 (Book review)
Bhattacharjee, Apurba K. 58 (Research communication)
Bhattacharjee, Somendra M. 816 (General article)
Bhattacharya, S. *see* Agarwala, B. K.
Bhattacharya, Samir 221 (Opinion)
Bhattacharyya, J. C. 486 (News); 715 (In this issue)
Bhattacharyya, S. P. *see* Maity, D. K.
Bilham, Roger 101 (General article)
Bisht, K. S. 496 (Scientific correspondence)
Biswas, Sukumar 719 (Research news); 721 (General article)
Bloom, David E. 822 (Introduction, special section); 865 (Article, special section)
Borah, Bijoylakshmi *see* Roy, Debjani
Bose, D. N. 474 (Personal news)
Brahmachary, R. L. *see* Sen, Srabani
Brown, Tim 840 (Article, special section)

Chakravarty, Charusita 739 (Review article)
Chalissery, Mary S. *see* Sane, R. T.
Chamyal, L. S. *see* Maurya, D. M.
Chander, Ramesh 223 (Commentary); 781 (Book review)
Chandra, Shaila 400 (Scientific correspondence); *see* Singh, Kamal Jeet
Chandran, Suresh 85 (Correspondence)
Chandrasekhar, J. 715 (In this issue)
Chandrasekhar, S. 557 (Historical commentary and notes)
Chandrashekara, K. 406 (Commentary); 470 (Book review)
Chandrashekaran, M. K. 895 (Correspondence)
Channing, C. *see* Winskill, L.
Chatterjee, Indu B. 747 (Review article)
Chaturvedi, U. C. *see* Mukerjee, Ruma
Chaudhuri, Amitabha 9 (Research news)
Chaurasia, S. S. 698 (Research communication)
Chidambaram, R. 13 (Commentary); 572 (General article)
Chin, James 828 (Article, special section)
Chopra, K. L. 565 (Correspondence)
Cubasch, U. *see* Lal, M.

Daniels, R. J. Ranjit 72; 703 (Book reviews); 569 (Scientific correspondence)
Das, Jyotirmoy *see* Dutta, Chitra
Das, K. K. *see* Roy, P. S.
Das, Manoj K. 946 (Book review); *see* Bal, Madhusmita; *see* Beuria, M. K.
Das, R. K. 96 (Scientific correspondence)
Das, Sahab *see* Shrivastav, Rohit
Das, S. K. 76 (Historical commentary and notes)
Das, Tapan Kanti 874 (Research communication)
Dash, B. C. *see* Das, R. K.
Desai, Umesh R. 893 (Correspondence)
De Sousa, S. N. *see* Dileep Kumar, M.
Dessai, A. G. *see* Karmalkar, N. R.
Dev, B. N. *see* Satyam, P. V.
Devasia, Preston *see* Gadgil, Madhav
Devi, M. *see* Nagabhushanam, R.
de Zoysa, Nandrani S. 857 (Article, special section)
Dhar, Deepak 368 (Book review)
Dileep Kumar, M. 672 (Review article)
Duraiswami, R. A. *see* Karmalkar, N. R.
Dutt, C. B. S. *see* Badarinath, K. V. S.
Dutta, Chitra 814 (Scientific correspondence)
Dutta, R. N. *see* Kar, R.
Dutta Roy, S. C. 6, 957 (Correspondence)
Dwivedi, G. N. 634 (Scientific correspondence)
Dwivedi, S. K. *see* Rao, V. V.
Dwyer, John M. 823 (Article, special section)

Esaki, L. 240 (General article)

Fingerman, M. *see* Nagabhushanam, R.

Gadagkar, Raghavendra 211 (In this issue); *see* Chandrashekara, K. (470)
Gadgil, Madhav 637 (Commentary)
Gadgil, Sulochana 649 (General article)
Gadre, Shridhar R. *see* Bhattacharjee, Apurba K.
Gahalaut, V. K. *see* Chander, Ramesh (223)
Ganapathi, N. G. 601 (Research communication)
Gandhi, K. S. 625 (Book review)
Ganesh Kumar, C. 237 (General article)
Ganeshaiah, K. N. 283, 563 (In this issue)
Gangal, S. V. 277 (Book review)
Ganguly, Ashok S. 193 (Generalia)
Garg, Rahul 1012 (Research communication)
Garg, V. K. *see* Rao, A. Bhaskara
Gaur, V. K. 83 (In this issue)
Geevan, C. P. 906 (General article)
George, M. D. *see* Dileep Kumar, M.
Gharai, Biswadip *see* Badarinath, K. V. S.
Ghose, S. K. *see* Satyam, P. V.
Ghosh, Mrinal K. *see* Chatterjee, Indu B.
Ghosh, Sabyasachi 937 (Research communication)
Giri, C. C. 458 (Research communication)
Giridharan, N. 493 (Scientific correspondence)
Glied, Sherry 849 (Article, special section)
Goodall, Jane 301 (Article, special section)
Gosal, S. S. *see* Bains, N. S.
Goswami, J. N. 1008 (Research communication); *see* Biswas, Sukumar (721)
Govindarajulu, V. 29 (General article)
Goyal, Navin *see* Maithani, P. B.
Gupta, P. *see* Chaurasia, S. S.
Gupta, S. K. 227 (Scientific correspondence)
Gururaja, M. N. *see* Patil, R. S.

Haque, M. R. *see* Sarker, A. A.
Hasan, Gaiti 626 (Book review)
Hegde, B. M. 946 (Book review)
Hegde, M. S. *see* Manoharan, S. Sundar
Hossain, M. *see* Islam, R.
Hosur, R. V. *see* Radha, P. K.

Islam, R. 494 (Scientific correspondence)
Iyengar, M. A. R. *see* Rajan, M. P.

Jagannathan, N. R. 448 (Research communication)
Jagetia, Ganesh Chandra *see* Ganapathi, N. G.
Jain, Abhineet *see* Roy, P. S.
Jain, K. P. *see* Garg, Rahul
Jain, M. C. *see* Ghosh, Sabyasachi
Jairajpuri, M. Shamim 370 (Book review)
Jaiswal, S. *see* Joshi, V. K.
Jayakumar, D. A. *see* Dileep Kumar, M.
Jayannavar, A. M. *see* Kumar, N. (57)
Jayaraman, A. 44 (Research article)
Jeyaseelan, A. T. *see* Badarinath, K. V. S.
Jha, Awadesh N. 1005 (Research communication)
Jha, C. S. *see* Badarinath, K. V. S.
Jha, Sumita 971 (Scientific correspondence)
Jha, Timir B. *see* Jha, Sumita
Jitendra Kumar *see* Kumar, Vandana A.
Joarder, O. I. *see* Islam, R.
John, S. 633 (Correspondence)
Johnston, Simon 521 (Review article)
Joseph, Ancy *see* Kumaran, K. P. N.
Joshi, N. V. 563, 631 (In this issue)
Joshi, V. K. 263 (Research communication)
Joshi, Y. P. 395 (Correspondence)

Kailas, S. *see* Chidambaram, R. (572)
Kalaivani, R. *see* Santhanalakshmi, J.
Kale, V. M. *see* Srivastava, A. K.
Kamble, C. V. *see* Kumaran, K. P. N.
Kanishkan, B. *see* Patil, R. S.
Kar, A. *see* Chaurasia, S. S.
Kar, R. 67 (Research communication)
Karim, M. R. *see* Islam, R.
Karmalkar, N. R. 767 (Research communication)
Kartha, C. C. 393 (News)
Kasturirangan, K. 732 (Review article)
Katti, Madhusudan 305 (Article, special section)
Kaul, V. K. *see* Parthasarathi, Ashok
Kaur, Ravinder *see* Bedi, J. S.
Kavitha, C. S. 957 (Correspondence)
Kaw, P. K. *see* Sen, Abhijit
Khandelwal, D. P. 88 (Correspondence)
Khanuja, Varjeet S. *see* Ajmani, Ranjeet S.
Khare, Avinash *see* Bhattacharjee, Somendra M.; *see* Rao, K. Srinivasa
Khattri, K. N. 161 (Research article); 361 (Research communication)
Khoshoo, T. N. 14 (Commentary)
Koteeswaran, R. 545 (Research communication)
Kotru, P. N. *see* Bhat, Sushma
Krishnaswami, S. *see* Trivedi, J. R.
Krishnaswamy, S. 182 (Research communication)
Kuber, V. V. *see* Sane, R. T.
Kulshrestha, S. M. 39 (General article)
Kumar, B. 929 (Research communication)
Kumar, N. 57 (Research communication); 492 (Research news)
Kumar, R. Senthil *see* Arumugam, M.
Kumar, Vandana A. 543 (Research communication)
Kumaraguru, P. *see* Patil, R. S.
Kumaran, K. P. N. 1023 (Research communication)
Kuri, G. *see* Satyam, P. V.

Lakhotia, S. C. 219 (Research news); 627 (Book review); 797 (News); *see* Singh, Bhupendra N.
Lal, M. 752 (Research article)
Lal, Shiv 833 (Article, special section)
Leblanc, M. *see* Karmalkar, N. R.

Mackay, Alan L. 151 (Research article)
Madan, Anup *see* Radha, P. K.
Madhavan, D *see* Parthasarathi, Ashok
Madhusudhanan, K. P. *see* Sivakumar, R.
Mahal, Ajay *see* Bloom, David E. (865)
Maheshwari, Ramesh 401 (Commentary)
Maithani, P. B. 270 (Research communication)
Maiti, P. K. *see* Chaurasia, S. S.
Maitra, Uday 964 (Research news)
Maity, D. K. 343 (Research article)
Majumdar, D. *see* Maity, D. K.
Majumder, S. K. *see* Kar, R.
Malthai, Jyotsna *see* Pandiyan, N.
Manchanda, Rohit 140 (Review article)
Manoharan, S. Sundar 336 (Research communication)
Maruthi Mohan, P. *see* Akthar, Md. Nam *see*
Masani, P. R. 288, 289 (Scientific correspondence)
Mashelkar, R. A. 317 (General article)
Mathur, Sanjay Kumar *see* Shrivastav, Rohit
Maurya, D. M. 610 (Research communication)

Mazumdar, Shyamalava *see* Das, Tapan Kanti
Mehrotra, K. N. *see* A. P. Padma Kumari
Menon, N. R. Madhava 881 (Book review)
Menon, S. *see* Sane, R. T.
Merh, S. S. *see* Maurya, D. M.
Mishra, S. *see* Goswami, J. N.
Misra, C. M. 17 (Scientific correspondence)
Mitra, A. N. 467 (Book review); 783 (Generalia); 802 (Commentary)
Mittal, P. K. *see* Adak, T.
Mohan Ram, H. Y. *see* Nair, M. N. B.
Mohanakumar, S. *see* Reghunadhan, K.
Moorthy, P. S. R. *see* Pandiyan, N.
Mukerjee, Ruma 900 (Scientific correspondence)
Mukhopadhyay, Chinmay K. *see* Chatterjee, Indu B.
Mukhopadhyay, Dhruba 189 (Book review)
Mukhopadhyay, Kakoli 965 (Research news)
Mukunda, N. 375 (Historical commentary and notes); 882 (Book review)
Muniyappa, K. 550 (Book review)
Muniyappa, V. 1031 (Book review)
Munuswamy, N. *see* Shanmugasundaram, G. K.
Muralikrishna, N. 718 (Scientific correspondence)
Murthy, G. S. *see* Venkatesh, N.
Mythili, J. B. *see* Thomas, Pious

Nadgauda, R. S. 277 (Book review)
Nagabhushanam, R 659 (Review article)
Nagaraja, G. *see* Venkatesh, N.
Nagarajan, K. 373 (Personal news)
Naik, P. C. 705 (Historical commentary and notes)
Nair, A. G. Ramachandran *see* Sivakumar, R.
Nair, M. N. B. 809 (Scientific correspondence)
Nampoori, V. P. N. *see* Tatavarti, Rao
Naqvi, S. W. A. *see* Dileep Kumar, M.
Narasimhan, N. S. 287 (Correspondence)
Narasimhan, R. 292 (Scientific correspondence); 701 (Book review)
Narayanaswamy *see* Ravindra Kumar, G. R.
Naren, Anitha 777 (Research communication)
Narlikar, Jayant 192 (Personal news); 969 (Scientific correspondence)
Narvekar, P. V. *see* Dileep Kumar, M.
Naseema, A. 461 (Research communication)
Nataraj, D. V. 434 (Research article)
Natarajan, R. *see* Rao, V. Purnachandra
Nautiyal, A. R. *see* Uniyal, R. C.
Nayar, M. P. 971 (Scientific correspondence)
Nene, Y. L. 191 (Book review)
Nityananda, R 483 (In this issue); 554 (Personal news)

Ojha, Devendra 518 (General article)
Ojha, J. 99 (Scientific correspondence)

Padayatti, Pius S. *see* Naseema, A.
Padhy, L. C. *see* Radha, P. K.
Padma, K. *see* Reddy, P. Chandrasekhara
Padma Kumari, A. P. 464 (Research communication)
Padmini, V. N. *see* Kasturirangan, K.
Pande, Kanchan *see* Trivedi, J. R.
Pandey, K. K. *see* Muralikrishna, N.
Pandian, T. J. *see* Koteeswaran, R.
Pandit, M. K. *see* Banerji, Subrata
Pandiyan, N. 899 (Scientific correspondence)
Pandiyan, Radha *see* Pandiyan, N.
Parnaik, Veena K. 71 (Book review)
Parthasarathi, Ashok 974 (General article)
Parthasarathy, K. R. 904 (General article)
Parthasarathy, R. 69 (Book review)
Parthasarathy, S. 529, 871 (Research communications)
Parthiban, S. 793 (News)
Patel, B. A. *see* Vyas, R. V.
Patel, D. J. *see* Vyas, R. V.
Patel, H. R. *see* Vyas, R. V.
Pati, Atanu Kumar 87 (Correspondence)
Patil, G. V. *see* Srivastava, A. K.
Patil, R. S. 97, 810 (Scientific correspondence)
Patil, Veeranagouda *see* Reddy, P. Chandrasekhara
Paulose, C. S. *see* Naseema, A.
Periasamy, Mariappan 287 (Correspondence)
Peterson, George *see* Sarda, Rabin M.
Phokela, A. *see* A. P. Padma Kumari
Prakash, Satya *see* Shrivastav, Rohit
Prasad, N. L. *see* Kasturirangan, K.
Prasad, T. G. *see* Naren, Anitha; *see* Reddy, P. Chandrasekhara
Prasad, V. *see* Manoharan, S. Sundar
Prasanna, B. M. 5 (Correspondence)
Pundlik, Savita S. *see* Bhattacharjee, Apurba K.
Puniyani, R. R. *see* Ajmani, Ranjeet S.

Rab, M. A. *see* Sarker, A. A.
Radha, P. K. 243 (Research account)
Radhakrishna, B. P. 891 (In this issue)
Radhakrishnan, P. *see* Tatavarti, Rao
Radhakrishnan, T. P. 217 (Research news)
Radhakrishnan, V. 189 (Book review)
Raghavendra, K. *see* Adak, T.
Raghunathan, P. *see* Jagannathan, N. R.
Rahman, A. 552 (Personal news)
Rahman, Abdul *see* Sharma, Kewal K.
Rajadnya, N. *see* Roy, P. S.
Rajagopalan, S. R. 3, 483 (In this issue)
Rajan, M. P. 592 (Research article)
Rajasree, K. *see* Tatavarti, Rao
Raju, K. K. V. S. *see* Ramesh Kumar, P. V.
Raju, K. S. *see* Bhat, Sushma
Ramachandran, S. *see* Maithani, P. B. *see* Muralikrishna, N.
Ramachandran, T. V. *see* Rajan, M. P.
Ramakrishnan, C. *see* Nataraj, D. V.
Ramani, A. V. 955 (In this issue)
Ramanna, Raja 897 (Correspondence)
Ramanujam, C. G. K. *see* Patil, R. S.
Ramasamy, Sm. 811 (Scientific correspondence)
Ramaseshan, S. 7, 301 (Editorial notes); 25 (General article); 893 (News)
Ramasubramanian, S. 231 (General article)
Ramesh Kumar, P. V. 763 (Research communication)
Ramya, K. *see* Parthasarathy, S.
Rangarajan, S. 410 (General article)
Rao, A. Bhaskara 958 (Correspondence)
Rao, A. P. Kamalakara *see* Senthil Kumar, P.
Rao, B. R. *see* Rao, V. Purnachandra
Rao, B. S. Ganga *see* Ramesh Kumar, P. V.
Rao, Ch. M. *see* Rao, V. Purnachandra
Rao, C. N. R. 216 (Research news)
Rao, G. U. 351 (Research article)
Rao, J. R. Lakshmana 896 (Correspondence)
Rao, K. Jayaraj *see* Ganesh Kumar, C.
Rao, K. Srinivasa 566 (News)
Rao, K. Vasudeva 397 (Scientific correspondence)
Rao, M. Rama *see* Patil, R. S.
Rao, Pitta Govinda 476 (Historical commentary and notes)

Rao, P. V. S. 420 (General article)
Rao, R. Balaji 369 (Book review)
Rao, R. R. 968 (Commentary)
Rao, U. R. *see* Kasturirangan, K.
Rao, V. Purnachandra 1017 (Research communication)
Rao, V. V. 932 (Research communication)
Rath, Satyajit 945 (Book review)
Ravan, Shirish A. *see* Roy, P. S.
Ravi *see* Bains, N. S.
Ravindra Kumar, G. R. 941 (Research communication)
Raviraja, N. S. *see* Sridhar, K. R.
Ray, Prasanta Kumar 549 (Book review)
Ray, S. L. *see* Goswami, J. N.
Reddy, P. Chandrasekhara 495 (Scientific correspondence)
Reddy, P. S. *see* Nagabhushanam, R.
Reghunadh, K. 456 (Research communication)
Rout, B. *see* Satyam, P. V.
Roy, Debjani 265 (Research communication)
Roy, P. S. 685 (Research article)

Saha, A. K. *see* Goswami, J. N.
Sajan, K. *see* Reghunadh, K.
Sandhu, D. K. *see* Joshi, V. K.
Sane, R. T. 899 (Scientific correspondence)
Santhanalakshmi, J. 918 (Research article)
Sarda, Rabin M. 835 (Article, special section)
Sarin, M. M. *see* Trivedi, J. R.
Sarker, A. A. 185 (Research communication)
Sarma, Amarendra *see* Roy, Debjani
Sarma, K. V. L. 397 (Scientific correspondence); 959 (News)
Sarojini, R. *see* Nagabhushanam, R.
Sashidhar, V. R. *see* Naren, Anitha
Satish, P. R. 336 (Review article)
Satpathy, L. *see* Naik, P. C.
Satyalakshmi, K. M. *see* Manoharan, S. Sundar
Satyam, P. V. 526 (Research communication)
Satyanarayana, P. *see* Srinivasan, G. J.
Savitha, J. 596 (Research article)
Sawhney, V. K. *see* Rao, G. U.
Saxena, R. K. *see* Misra, C. M.
Saxena, S. K. 914 (General article)
Sekar, Vaithilingam *see* Unnamalai, N.
Sen, Abhijit 73 (Personal news)
Sen, D. 426 (Review article)
Sen, Srabani 895 (Correspondence)
Sengupta, S. *see* Das, S. K. *see* Sen, D.
Senthil Kumar, P. 258 (Research communication)
Shankar, P. N. 483 (In this issue)
Shankar, R. 86 (Correspondence)
Shanmugasundaram, G. K. 229 (Scientific correspondence)
Shanmugasundaram, S. *see* Krishnaswamy, S.
Sharma, Anand K. *see* Virk, H. S.
Sharma, Kewal K. 613 (Research communication)
Sharma, M. K. 619 (Research communication)
Sharma, P. *see* Gupta, S. K.
Sharma, Ram S. 1031 (Book review)
Sharma, S. K. *see* Jayaraman, A.
Sharma, S. R. *see* Rao, V. V.
Sharma, V. P. *see* Adak, T.
Sheela, S. G. *see* Koteeswaran, R.
Shivanna, K. R. *see* Nair, M. N. B.; *see* Rao, G. U.
Shrivastav, M. M. *see* Shrivastav, Rohit
Shrivastav, Rohit 272 (Research communication)
Shrivastav, Shobhit *see* Shrivastav, Rohit
Shrivastava, P. K. *see* Muralikrishna, N.
Simlot, M. M. *see* Sharma, M. K.
Singh, Bagwan D. 998 (Research article)
Singh, B. N. 771 (Research communication)
Singh, Jaswinder *see* Bains, N. S.
Singh, J. S. 293 (Scientific correspondence)
Singh, Kamal Jeet 228 (Scientific correspondence); *see* Chandra, Shaila
Singh, Lambodar P. 801 (Correspondence)
Singh, N. Bhupendra 178 (Research communication)
Singh, Rajendra *see* Maithani, P. B.; *see* Muralikrishna, N.
Singh, Ramesh P. 286 (News); *see* Aman, Ashutosh
Singh, Sarnam *see* Roy, P. S.
Singh, Smita *see* Singh, J. S.
Singh, Umesh K. *see* Aman, Ashutosh
Sinha, Anindya 296 (Article, special section); 903 (Research news)
Sinha, Suresh K. *see* Ghosh, Sabyasachi
Sisodia, Seema *see* Singh, B. N.
Sitaramam, V. 89 (Opinion)
Sittitrai, Werasit *see* Brown, Tim
Sivakumar, R. 23 (Scientific correspondence)
Sokhi, S. S. *see* Bedi, J. S.
Soman, K. *see* Kumaran, K. P. N.
Sorkhabi, Rasoul B. 489 (News)
Sowdhamini, R. *see* Nataraj, D. V.
Sreenivasan, P. 617 (Research communication)
Sridhar, K. R. 570 (Scientific correspondence)
Srinivas, V. 891 (In this issue)
Srinivasan, D. 370 (Book review)
Srinivasan, G. J. 534 (Research communication)
Srinivasan, K. 577 (General article)
Srinivasan, N. *see* Nataraj, D. V.
Srivastava, A. K. 397 (Scientific correspondence)
Srivastava, P. N. 288 (Correspondence)
Subbarao, Sarala K. *see* Adak, T.
Subramanian, C. V. 986 (General article); *see* Savitha, J.
Subramanyam, S. V. *see* Manoharan, S. Sundar
Subramoniam, A. 898 (Scientific correspondence)
Sundaravel, B. *see* Satyam, P. V.
Swaminathan, M. S. 276, 471 (Book reviews)
Swarup, D. *see* Rao, V. V.
Swathi, P. S. 285 (News)

Talwar, G. P. 473 (Personal news)
Tarafdar, J. C. 541 (Research communication)
Tatavarti, Rao 678 (Research article)
Tauro, P. 717 (Correspondence)
Tewari, V. C. *see* Kumar, B.
Thakur, B. B. *see* Lal, Shiv
Thamban, M. *see* Rao, V. Purnachandra
Thomas, Pious 94 (Scientific correspondence)
Tiwari, S. C. 213 (Correspondence)
Tripathi, A. K. *see* Dwivedi, G. N.
Trivedi, J. R. 171 (Research communication)
Tyagi, R. C. *see* Dwivedi, G. N.

Uberoi, Chanchal 74 (Personal news)
Udaya Kumar, M. *see* Naren, Anitha; *see* Reddy, P. Chandrasekhara
Udgaonkar, B. M. 197 (Historical commentary and notes)
Uma Devi, P. *see* Bisht, K. S.
Uma Shaanker, R. 955 (In this issue)
Uniyal, R. C. 358 (Research communication)
Unnamalai, N. 939 (Research communication)
Unnikrishnan, C. S. 896 (Correspondence)
Upadhyay, R. R. 717 (Correspondence)

Vallabhan, C. P. G. *see* Tatavarti, Rao
Varalakshmi, B. *see* Pandiyan, N.
Vasan, S. S. *see* Arya, S. K.
Vatsala, T. M. *see* Senthil Kumar, P.
Vencatesan, Jayashree *see* Daniels, R. J. Ranjit (569)
Venkataraman, Arun B. 934 (Research communication)
Venkatesan, V. K. *see* Parthasarathy, S. (871)
Venkatesh, N. 48 (Research article)
Vidyala, V. *see* Tatavarti, Rao
Vijayan, E. *see* Sreenivasan, P.
Virdi, J. S. 491 (Correspondence); 624, 1032 (Book reviews)
Virk, H. S. 452 (Research communication)
Vishwakarma, R. G. *see* Abdussattar
Vishwakarma, Rajesh K. 62 (Research communication)
Voss, R. *see* Lal, M.
Vyas, R. V. 635 (Scientific correspondence)

Walia, Vivek *see* Virk, H. S.
Wang, S. Y. *see* Jayaraman, A.
Waran, N. K. 303, 308 (Articles, special section); *see* Winskill, L.
Waszkewitz, J. *see* Lal, M.
Wiedenbeck,M. *see* Goswami, J. N.
Winskill, L. 310 (Article, special section)

Yadav, M. L. *see* Dwivedi, G. N.
Yogananda, C. S. 702 (Book review)
Young, R. *see* Winskill, L.

SUBJECT INDEX

Articles in Commentary, Generalia, Historical commentary and notes, News, Opinion, Research account, Research news; Books reviewed and letters in Correspondence and Scientific correspondence are also listed alphabetically under the respective category name.

Acacia auriculiformis 495 (Scientific correspondence)
Academy document on science education 88 (Correspondence); 89 (Opinion)
Academy Fellowship 969 (Scientific correspondence)
Accelerators 572 (General article)
Acid phosphatase 541 (Research communication)
Active gravitational mass 924 (Research communication)
Adaptive features 587 (Review article)
Aegle marmelos 494 (Scientific correspondence); 932 (Research communication)
Aggressive behaviour 491 (Research news)
AIDS 822 (Introduction, special section); 823, 828, 833, 835, 840, 849, 865 (Articles, special section)
AIDS and tuberculosis 833 (Article, special section)
AIDS, economic analysis 865 (Article, special section)
AIDS, lessons from USA 849 (Article, special section)
Akkulam lake 454 (Research communication)
Albizia lebbek 358 (Research communication)
Alfven, Hannes 73, 74 (Personal news)
Aliphatic-aromatic balance 67 (Research communication)
Alpha-logger 452 (Research communication)
α-turn 434 (Research article)
Ancient gold mining sites 634 (Scientific correspondence)
Animal rights 296 (Article, special section)
Animal welfare 283 (In this issue)
Anomalous cosmic rays 721 (General article)
Anorthosites 21 (Scientific correspondence)
Antihyperglycaemic effects 932 (Research communication)
Anuran acoustic communication 265 (Research communication)
Apple pomace 263 (Research communication)
Applied research 12 (Commentary)
Archaean crust 613 (Research communication)
Armillaria millea 986 (General article)
Artificial satellites 732 (Review article)
Ashtekar program 499 (General article)
Aspergillus flavipes 596 (Research article)
Atmospheric deposition 272 (Research communication)
Autoregulation 178 (Research communication)
Avicennia marina 926 (Research communication)
Axenic culture 21 (Scientific correspondence)
Azolla sp. 97 (Scientific correspondence)

Bacillus sphaericus 695 (Research communication)
Bacillus thuringiensis 939 (Research communication)
Bacterial pathogenesis 1032 (Book review)
Bacterial virulence 965 (Research news)
Bairam–Belkher area 397 (Scientific correspondence)
Basic research 12 (Commentary)
Big bang cosmology 924 (Research communication)
Bilateral outgrowths 771 (Research communication)
Binary pulsar 521 (Review article)
Bioacoustic analysis 265 (Research communication)
Biodiversity 14 (Commentary); 906 (General article)
Biogas inhibition 619 (Research communication)
Biological control 939 (Research communication)
Biological survey 14 (Commentary)
Bioremediation 1028 (Research communication)
Blood transfusion 857 (Article, special section)
Blood viscosity 877 (Research communication)
Bombay offshore 17 (Scientific correspondence)
Bombyx mori 185 (Research communication)

Book reviews
The Action Principle in Physics 368
Annual Review of Biochemistry, Vol. 63, 1994 550
Annual Review of Cell Biology, Vol. 10, 1994 71
Annual Review of Immunology, Vol 13, 1995 945
Annual Review of Medicine, Vol. 46, 1995 946
Annual Review of Neuroscience, Vol. 18 , 1995 626
Annual Review of Phytopathology, Vol. 32, 1994 191
Bacterial Pathogenesis of Plants and Animals: Molecular and Cellular Mechanisms 1032
Beyond the Quantum Paradox 367
Botany in India: History and Progress 371
The Collected Papers of Albert Einstein: Vol. 5 – *The Swiss Years: Correspondence, 1902–1914* 882
Contributions to Operator Theory and its Applications 69
Crustal Evolution of Singhbhum – North Orissa, Eastern India 189
Deep Continental Structure of India: A Review 1031
Development of Ideas in Physics 702
Environment and Adaptive Biology of Plants 72
Ethnobotany and the Search for New Drugs 548
Environmental Health Perspective: Supplements: Molecular Mechanisms of Metal Toxicity and Carcinogenicity 549
Functional Dynamics of Phytophagous Insects 470
The Guide to the Galaxy 189
In vitro Culture and its Applications in Horticulture 277
Laboratory Manual of Organic Chemistry 369
Law, Ethics and HIV: Proceedings of the UNDP Intercountry Consultation 881
Leks 703
Life in Moving Fluids 625
Mahatma Gandhi: An Apostle of Applied Human Ecology 276
A Modern Introduction to Ancient Indian Mathematics: Indian Mathematics and Astronomy 702
Molecular Genetics of Bacterial Pathogenesis 624
Ocean Technology – Perspectives 370
Perspectives in Nematode Physiology and Biochemistry 370
Quark and Jaguar 467
Recent Trends in Aerobiology, Allergy and Immunology 277
Rice Tungro 1994 1031
Science in Society: Some Perspectives 69
Speaking Minds: Interviews with Twenty Eminent Cognitive Scientists 701
The State of Food and Agriculture 1993 and 1994 471
Textbook of Molecular Biology 627
Tropical Diseases: Molecular Biology and Control Strategies 946
Uttarkashi Earthquake 781

Born–Oppenheimer description 715 (In this issue)

Bose–Einstein condensation 492 (Research news)
Bose, S. N. 76 (Historical commentary and notes)
Botryococcus 810 (Scientific correspondence)
Brassinolide 777 (Research communication)
Bundelkhand craton 613 (Research communication)

Calcium 777 (Research communication)
Calcutta earthquake of 1737 476 (Historical commentary and notes)
Carbon and oxygen isotopic records 691 (Research communication)
Carbon dioxide 672 (Review article)
Carbonyl addition reaction 343 (Research article)
Cauvery basin 97, 810 (Scientific correspondence)
CD_4 lymphocytes 823 (Article, special section)
CD_8 lymphocytes 823 (Article, special section)
CD-ROM 285 (News)
Cell elongation 777 (Research communication)
Cellular relevance 9 (Research news)
Central limit theorem 904 (General article)
Chaca chaca 99 (Scientific correspondence)
Chandra Sekhar, Samanta 705 (Historical commentary and notes)
Chandrasekhar, Subrahmanyan 372; 554 (Personal news)
Charnockites 941 (Research communication)
Chatterjee, S. D. 474 (Personal news)
Chemical Abstracts 285 (News)
Chemical structures 151 (Research article)
Chemotherapy 257 (Research communication)
Chlorofluorocarbons 793 (News)
Chromosome elimination 458 (Research communication)
Circadian rhythms 626 (Book review)
Climate change 39 (General article); 752 (Research article)
Climate change and agriculture 649 (General article)
Coal 592 (Research article)
Cockerel 698 (Research communication)
Codon usage 182 (Research communication)
Columbite-tantalite bearing pegmatite 718 (Scientific correspondence)

Commentary
Are herbaria redundant? 968
Census of India's biodiversity: Tasks ahead 14
Finding *patterns* in nature's maze: An endless quest 406
Implications of recent levelling observations for Tehri and other high dams in the Himalaya 223
Importance of basic research in the context of global change 12
India's scientific community: A cultural crisis? 802
Intellectual property rights and biological resources: Specifying geographical origins and prior knowledge of uses 637
Whither microbiology? 401

Commercial sex 840 (Article, special section)
Communication satellites 410 (General article)
Complementarity principle 426 (Review article)
Complex adaptive systems 467 (Book review)
Conservation 14 (Commentary)
Constituents of matter 397 (Scientific correspondence)
Contraceptive use 577 (General article)
Cordierite gneisses 763 (Research communication)

Correspondence
The Academy document: A perceptive appraisal of science education in India 87
The Academy document: A teacher's view 88
Assessment of research impact of engineering institutions 565
A crisis in Engineering education 957
Distortion of Indian degrees 6
Early research in India on primate tool using 895
Fancy and facts in science and art 5
Feast of the Gods? 717
A good beginning 491
If education is important, so must be teachers! 801
Karnatak music 896
Microbiology in India 717
More on microbiology 895
More on NSU 288
Necessity of broadening science base in India 287
NSU, Academy Document and science education: Reflections and an operation model 958
Nurturing young scientists 893
Patenting biotechnologies 85
Ph D students, research supervisors and university research: Some observations 633
Proposal for a National Science School 396
Revitalization of Indian science 213
Vitamin C: A vital nutrient 957
What NSU report leaves unanswered! 395
Why not platinum? 86

Cosmic rays 715 (In this issue)
Creative potential 5 (Correspondence)
Cretaceous-Tertiary boundary 1012 (Research communication)
Crude oil 596 (Research article)
Crustal growth 613 (Research communication)
Crutzen, Paul J. 793 (News)
Cryopreservation 899 (Scientific correspondence)
Cryptobiosis 229 (Scientific correspondence)
Culex quinquefasciatus 695 (Research communication)
Cultural crisis 802 (Commentary)
Cuon alpinus 934 (Research communication)
Cyanobacteria 182 (Research communication)
Cybernetic system 288 (Scientific correspondence)
Cyclones 39 (General article)
Cyclotron 572 (General article)
Cytochrome c oxidase 874 (Research communication)
Cytokine antagonism 900 (Scientfic correspondence)

Deep Continental structure 1031 (Book review)
Delphi survey method 828 (Article, special section)
Deltamethrin 464 (Research communication)
Density determination 534 (Research communication)
Detection limits 640 (General article)
Dholes 903 (Research news); 934 (Research communication)
Diabetes 461 (Research communication)
Dioxins 237 (General article)
Dibenzodioxins 237 (General article)
DNA-binding domain 243 (Research account)
Donor selection 857 (Article, special section)
Drag force constant 534 (Research communication)
Drosophila 797 (News)
Drosophila bipectinata 771 (Research communication)

Earth's core 914 (General article)
Earthquake 161 (Research article); 361, 452, 772 (Research communications)
Economics of AIDS 849 (Article, special section)
Einstein, Albert 882 (Book review)
Electron transfer 739 (Review article)
Electrophysiology 140 (Review article)
Embryonic tissues 494 (Scientific correspondence)
Embryo rescue 621 (Research communication)

Endoyphytes 570 (Scientific correspondence)
Engineering education 957 (Correspondence)
Engineering institutions 565 (Correspondence)
Enkephalins 659 (Review article)
Environmental design 308 (Article, special section)
Environmental health 549 (Book review)
Environmental pollution 592 (Research article)
Epididymal spermatozoa 899 (Scientific correspondence)
Epitopes 48 (Research article)
Eryx conicus 258 (Research communication)
Eryx johnii 258 (Research communication)
Ethephon 809 (Scientific correspondence)
Ethnobotany 548 (Book review)
Exact solution 816 (General article)
Examination system 87 (Correspondence)
Experimental physics 25 (General article)
Exploitation of animals 296 (Article, special section)

Fairy shrimp 229 (Scientific correspondence)
Falling liquid drops 534 (Research communication)
Feature vector 420 (General article)
Female literacy 577 (General article)
Female reciprocal call 265 (Research communication)
Fertility 577 (General article)
Fertility patterns 563 (In this issue)
Flavonoids 23 (Scientific correspondence)
Flexicrystallography 151 (Research article)
Flowering plants 971 (Scientific correspondence)
Fluorescing macerals 998 (Research article)
Foetal exposure 496 (Scientific correspondence)
Food contaminants 237 (General article)
Fowler, Willy 192 (Personal news)
Fractal description 361 (Research communication)
Functional imaging 391 (In this issue)
Functional magnetic resonance imaging (FMRI) 448 (Research communication)
Fungal biosorbent 1028 (Research communication)

Galaxy kinematics and dynamics 518 (General article)
Galaxy stellar content 518 (General article)
Gamma ray bursts 732 (Review article)

Generalia
In the pursuit of science: The 'subjective' element 783
Who should look at stars 193

Genetic markers 926 (Research communication)
Genetic resources 637 (Commentary)
Genotoxicity 96 (Scientific correspondence)
Geochemical monitoring 272 (Research communication)
Geotechnical implications 454 (Research communication)
Giant magnetoresistance 356 (Research communication)
GnRH 617 (Research communication)
Gondwana flora 397 (Scientific correspondence)
Goral habitat 685 (Research article)
Govindachari, T. R. 373 (Personal news)
Great Himalayan earthquakes 279 (Historical commentary and notes)
Greenhouse gases 293 (Scientific correspondence); 672 (Review article); 752 (Research article)
Ground support systems 410 (General article)
Gudakhu 96 (Scientific correspondence)
Gujarat plain 610 (Research communication)

Habitat suitability analysis 685 (Research article)
Harzburgites 767 (Research communication)
Heat shock response 178 (Research communication)
Heat tolerane 351 (Research article)
Heavy elements 295 (Scientific correspondence)
Heavy metals 272 (Research commnication)
Helichrysum bracteatum 23 (Scientific correspondence)
Helicoverpa armigera 464 (Research communication)
Hemp 893 (News)
Hepatoprotection 898 (Scientific correspondence)
Hepatoprotective drugs 898 (Scientific correspondence)
Herbaria 968 (Commentary)
Heterosexual behaviour 840 (Article, special section)
Higher education 213 (Correspondence)
Himalaya–Karakoram–Tibet region 489 (News)
Himalayan earthquakes 101 (General article)
Himalayan geodesy 223 (Commentary)
Historic seismicity 279 (Historical commentary and notes)

Historical commentary and notes
Eugene Paul Wigner – A tribute 375
On Ramanujan's bust 557
A probe into the Calcutta earthquake of 1737 476
Professor S. N. Bose's contribution to quantum statistics – A critical review 76
Samanta Chandra Sekhar: Life and work 705
Scientific tradition and other traditions 197
Two undocumented historic earthquakes and additional information on the major earthquake of 1 September 1803 from north India 279

History and philosophy 197 (Historical commentary and notes)
HIV 822 (Introduction, special section); 823, 828, 835, 840, 857, 861, 865 (Articles, special section); 870 (Book review)
Hornblende mineral reference sample 640 (General article)
Human-animal interactions 303 (Article, special section)
Human chorionic gonadotropin 48 (Research article)
Human filariasis 538 (Research communication)
Human interactions 783 (Generalia)
Humic acids 67 (Research communication)
Hydrogen bonds 434 (Research article)
Hydrogen electrode 871 (Research communication)
Hyoscyamus muticus 458 (Research communication)
Hypoglycaemic effect 932 (Research communication)

Igneous cryatallization 21 (Scientific correspondence)
Immunocytochemistry 178 (Research communication)
Indian degrees 6 (Correspondence)
Indian food practices 415 (General article)
Indian Gondwana coals 998 (Research article)
Indian ocean 672 (Review article)
Indian population 563 (In this issue)
Indian rivers 171 (Research communication)
Indian science 197 (Historical commentary and notes); 213 (Correspondence)
Indian universities 633,801 (Correspondence)
Indus ophiolite belt 767 (Research communication)
Information network 906 (General article)
Inhibitory glycine receptor 336 (Review article)
Intellectual property rights 29 (General article); 637 (Commentary)
Internalwaves 678 (Research article)
International collaboration 12 (Commentary)
International Lithosphere Project 286 (News)
Interstellar medium 721 (General article)
Inter-University Centre for Astronomy and Astrophysics (IUCAA) 25 (General article)
Invertebrates 659 (Review article)

In vitro seeds 543 (Research communication)
Ion implantation 526 (Research communication)
Ion-microprobe 1008 (Research communication)
Ionic surfactant 874 (Research communication)
Ionization states 721 (General article)
Ising model 816 (General article)
ISI's electronic library 485 (News)
Isotope trends 929 (Research communication)
Isovaleric acids 619 (Research communication)

Journal editing 393 (News)
Junction potentials 140 (Review article)

Kalinga Prize 215 (News)
Kalpasar 227 (Scientific correspondence)
Karaya gum 809 (Scientific correspondence)
Karnatak music 896 (Correspondence)
Kramers' theory 129 (Review article)

Laser ablation 640 (General article)
Laser beam deflection 678 (Research article)
Laser-heating thermodynamics 914 (General article)
Laser spectroscopy 129 (Review article)
LDH isoenzymes 258 (Research communication)
Lead 698 (Research communication)
Leopards 903 (Research news)
Leks 703 (Book review)
Leptons 959 (News)
Leucostoma 939 (Research communication)
Lewis, Edward B. 797 (News)
Lignite 592 (Research article)
Lipid enzymes 596 (Research article)
Lipid peroxidation 601 (Research communication)
Liquid junction 529 (Research communication)
Lithium 252 (Research article)
Liv.52 601 (Research communication)
Low serum cholesterol 491 (Research news)
Lymphokines 823 (Article, special section)

Mahatma Gandhi 276 (Book review)
Maize 621 (Research communication)
Male fructification 400 (Scientific correspondence)
Mangroves 926 (Research communication)
Marker processes 231 (General article)
Matter-wave interferrometry 426 (Review article)
Maturation 358 (Research communication)
Meloidogyne incognita 635 (Scientific correspondence)
Metallic diffusivity 57 (Research communication)
Metallogeny 252 (Research article)
Metal removal 1028 (Research communication)
Methane efflux 293 (Scientific correspondence)
Methane production 937 (Research communication)
Micellar catalysis 918 (Research article)
Microbiology 401 (Commentary); 895 (Correspondence)
Microbiology in India 717 (Correspondence)
Microbiological research 401 (Commentary)
Microcirculation 877 (Research communication)
Microemulsion media 918 (Research article)
Microprobe geochemistry 1017 (Research communication)
Microquantity liquids 534 (Research communication)
Microvasculature 99 (Scientific correspondence)
Micula prinsii 1012 (Research communication)
Milky way structure 518 (General article)
Modern science and technology 29 (General article)
Molecular electrostatic potentials 58 (Research communication)
Molecular genetics 624 (Book review)
Molecular imprinting 317 (General article)
Molina, Mario J. 793 (News)
Monoclonal antibodies 48 (Research article)
Monodeiodinase activity 698 (Research communication)
Morarji Desai and rhesus macaque 296 (Article, special section)
Mosquito control 695 (Research communication)
Motor cortex 448 (Research communication)
Moving fluids 625 (Book review)
Mulberry 67,185 (Research communications)
Multiple shoot production 495 (Scientific correspondence)
Muscardine fungi 635 (Scientific correspondence)
Mushrooms 986 (General article)
Mutagens 58 (Research communication)
Myb-DNA system 243 (Research account)

Nanoplankton marker 1012 (Research communication)
Nanotechnology 317 (General article)
Naphthyl phosphate 541 (Research communication)
National Science School 396 (Correspondence)
National Science University 287, 395, 958 (Correspondence)
Natural resources 569 (Scientific correspondence)
Needham, Joseph 552 (Personal news)
Nematode physiology 370 (Book review)
Neodymium heptamolybdate 607 (Research communication)
Neotectonics 811 (Scientific correspondence)
1833 Nepal earthquake 101 (General article)
Neuropeptides 617 (Research communication)
Neurotransmitters 617 (Research communication)
Neutron stars 732 (Review article)
New radiation belt 719 (Research news)

News
Amending the Patents Act – An appeal from the National Working Group on Patent Laws 7
Chemical Abstracts on CD-ROM 285
Hemp in the manufacture of paper 893
10th Himalaya–Karakoram–Tibet Workshop, Switzerland, 1995 – Conference report 489
Intensive course on 'applications of remotely sensed data on ocean colour' 285
International Lithosphere Project 286
ISI's electronic library 485
1994 Kalinga Prize 215
New appointments 215
Nobel leptons 959
Nobel Peace Prize for 1995 to Joseph Rotblat and the Pugwash Conferences on Science and World Affairs 717
The 1995 Nobel Prize in Medicine: A tribute to the power of formal genetics 797
Peer review under scrutiny 393
Pioneers of ozone study bag chemistry Nobel 793
Report on the 'International conference on symmetry methods in physics' (ICSMP-95) 566
Third rice genetics symposium 963
Total solar eclipse – October 1995 486

Nicotinic acetylcholine receptor 336 (Review article)
Niobium 270 (Research communication)
Nitrous oxide 672 (Review article)
NMR 211 (In this issue); 243 (Research account)
Nobel Peace Prize 717 (News)
Nobel Prize, Chemistry 793 (News)
Nobel Prize, Medicine 797 (News)
Nobel Prize, Physics 959 (News)
Nonlinear dynamics 317 (General article)
Nonlinear optics 217 (Research news)

Nuclear tracks 721 (General article)
Nucleotide-protein 814 (Scientific correspondence)
Nusslein-Volhard, Christianne 797 (News)

Ocean colour 285 (News)
Ocean technology 370 (Book review)
Onsager 816 (General article)
Onshore gas well 604 (Research communication)
Opal phytoliths 17 (Scientific correspondence)

Opinion
Declining phase of quality research in our universities 221
The Indian Academy of Sciences document on science education in universities: A criticism 89

Opioid peptides 631 (In this issue)
Opioids 617 (Research communication)
Oxidative damage 747 (Review article)
Oxide electronics 216 (Research news)
Oxide films 356 (Research communication)
Ozone layer 793 (News)

Palaeoecology 1023 (Research communication)
Palynoflora 1023 (Research communication)
Patenting biotechnologies 85 (Correspondence)
Patent Laws 7 (News)
Path integrals 739 (Review article)
Pathogenesis 965 (Research news)
Patterns 406 (Commentary)
Peer review 393 (News)
Periastron 521 (Review article)
Peripheral blood lymphocytes 1005 (Research communication)
Peripheral erythrocytes 96 (Scientific correspondence)
Perl, Martin L. 959 (News)
Peroxodisulphate reactions 918 (Research article)

Personal News
A doyen of Indian physicists (An obituary of S. D. Chatterjee) 474
Joseph Needham (1900–1995) (An obituary) 552
A pioneer in modern plasma physics (An obituary of Hannes Alfvén) 73
Prof. T. R. Govindachari – An appreciation 373
Professor Willy Fowler (1911–1995) (An obituary) 192
Remembering Hannes Alfvén and his work (An obituary of Hannes Alfvén) 74
Subrahmanyan Chandrasekhar (An obituary) 554
A visionary and courageous scientist (An obituary of Jonas Salk) 473

Peshwa period 279 (Historical commentary and notes)
Petunia 351 (Research article)
Phase diagram 914 (General article)
Ph D students 633 (Correspondence)
Phosphorite 1017 (Research communication)
Phosphorylation 336 (Review article)
Phytophagous insects 470 (Book review)
Planktonic foraminifera 691 (Research communication)
Plant cells 191 (Book review)
Platinum 86 (Correspondence)
Pod formation 543 (Research communication)
Pollen function 351 (Research article)
Postnatal growth 496 (Scientific correspondence)
Precambrian-Cambrian boundary 929 (Research communication)
Predatory ladybird beetles 587 (Review article)
Pressure-induced phase transition 44 (Research article)
Propensity of amino acid 434 (Research article)
Protein folding 9 (Research news)
Pseudofruit 94 (Scientific correspondence)
Public knowledge 637 (Commentary)
Puccinia romagnoliana 21 (Scientific correspondence)

Quantum chemical bond order 343 (Research article)
Quantum gravity 483 (In this issue); 499 (General article)
Quantum Monte Carlo 739 (Review article)
Quantum mechanics 367 (Book review)
Quantum statistics 76 (Historical commentary and notes)
Quantum universe 467 (Book review)
Quaternary basin 610 (Research communication)

Radiation dose 592 (Research article)
Radon 452 (Research communication)
Rainfall pattern 649 (General article)
Raman spectroscopy 3 (In this issue); 44 (Research article)
Ramanujan's bust 557 (Historical commentary and notes)
Random numbers 904 (General article)
Reactive oxygen species 747 (Review article)
Reines, Frederick J. 959 (News)
Relaxation phenomenon 129 (Review article)
Remote sensing 631 (In this issue); 685 (Research article); 811 (Scientific correspondence)
Reproductive behaviour 587 (Review article)

Research account
NMR studies on the Myb-DNA system 243

Research impact 565 (Correspondence)

Research news
Binding and release of polypeptides during groEl-mediated folding: Its cellular significance? 9
Bose–Einstein condensation in a dilute atomic vapour 492
Designer molecules: A homogenic sensor of creatinine 964
Discovery of the new radiation belt of the earth 719
Does low serum cholesterol increase suicide and aggressive behaviour? 491
Driven up the tree: Why do wild dogs harass leopards? 903
Noncentrosymmetric monolayers from centrosymmetric molecules – A fresh approach to materials for frequency doubling 217
Oxide electronics 216
Pathogenesis: Generating new paths in understanding bacterial virulence 965
Targeting RNA in the cell 219

Rhizosphere 541 (Research communication)
Rice 937 (Research communication); 963 (News)
Rice field 293 (Scientific correspondence)
Rice tungro 1031 (Book review)
RNA 219 (Research news)
Rotblat, Joseph 717 (News)
Roughness measurement 526 (Research communication)
Rowland, F. Sherwood 793 (News)

Salk, Jonas 473 (Personal news)
Sargipali sulphide deposit 62 (Research communication)
Satellite data 604 (Research communication)
Schrodinger equation 57 (Research communication)
Science education 87,491,801 (Correspondence)
Science in India 287 (Correspondence)

Scientific correspondence
The Academy fellowship problem 969
Angioarchitecture of the glomerulus of a swamp catfish, *Chaca chaca* (Ham.) (Chicidae, Siluriformes) 99
Azolla sp. from the Early Cretaceous of Cauvery basin, South India 97
Comments on 'What kind of a system is a human being? A cybernetic system?" (*Curr. Sci.*, 1994, **67**, 936–940) 288
Comments on the review of the monograph *The Scientific Methodology in the Light of Cybernetics* (*Curr. Sci.*, 1994, **67**, 954–956) 289
Cryopreservation of epididymal spermatozoa 899
Cytokine antagonism by active vaccination 900
Development of cultured tomato anther to a fruit-like structure accompanied by *in vitro* ripening 94
Effect of foetal exposure to low-dose X-rays on the postnatal growth of mouse 496
Endophytes – A crucial issue 570
Enechelon faults along West Coast of India and their geological significance 811
Ethephon enhances karaya gum yield and wound healing response: A preliminary report 809
A fast growing cell line of *Taxus baccata* L. (Himalayan yew) as a potential source of taxol precursor 971
First report of the Late Permian sporangiate fructification *Nesowalesia* Pant from India 400
Genotoxicity of 'gudakhu', a tobacco preparation used in Orissa 96
In-situ preserved *Vertebraria* axes in Ib river coalfield, Orissa 228
Interpretation of nucleotide/protein sequence data: Some pitfalls 814
In-vitro axillary bud break and multiple shoot production in *Acacia auriculiformis* by tissue culture technique 495
Massif-type anorthosite near Ayialur, Dindigul, Anna District, Tamil Nadu, South India 21
Methane emission from rice paddies: Need for a downward revision of global estimate 293
Nomenclature of super heavy elements 295
A note on the occurrence of Columbite–Tantalite-bearing pegmatite at Limboi, District Sabarkantha, Gujarat 718
Occurrence of *Botryococcus* from the Lower Cretaceous rocks of Cauvery basin, Tamil Nadu 810
Parasitism of muscardine fungi on *Meloidogyne incognita* eggs 635
Polyphenolic components of flowers of *Helichrysum bracteatum* 23
Recent finds of ancient gold mining sites in south Uttar Pradesh 634
Record of opal phytoliths and associated palynofossils from Indian crude oils 17
Record of Upper Gondwana flora from Bairam–Belkher area of district Amravati, Maharashtra and district Betul, Madhya Pradesh 397
Regeneration of *Aegle marmelos* (L). Corr. plants *in vitro* from callus cultures of embryonic tissues 494
Role of phospholipase C in the cryptobiotic cysts of the fairy shrimp *Streptocephalus dichotomus* 229
Storing river runoff in the Gulf of Khambhat: A critique of the 'Kalpasar' concept 227
Supplementary comments on 'Endemic, rare and threatened flowering plants of South India' 971
'Swiss' strain of rats? 493
A technique for axenic culture of *Puccinia romagnoliana* – A potential bioherbicide for purple nutsedge 21
Traditional ecological knowledge and sustainable use of natural resources 569
The ultimate constituents of matter 397
Too many hepatoprotective drugs and a little hepatoprotection 898

Scientific software 802 (Commentary)
Script recognition 420 (General article)
Seamless chemical engineering science 317 (General article)
Sea-surface temperature 649 (General article)
Second harmonic generation 217 (Research news)
Seed development 358 (Research communication)
Seismic hazard 83 (In this issue)
Semiconductor superlattice 240 (General article)
Seismic hazards 161 (Research article); 223 (Commentary)
Sepia mutants 771 (Research communication)
Setaria digitata 538 (Research communication)
Shoemaker Levy 9 193 (Generalia)
Siddhanta Darpana 705 (Historical commentary and notes)
Singhbhum 1008 (Research communication)
Single crystals 607 (Research communication)
Single-ion activities 483 (In this issue); 529 (Research communication)
Smoking 877 (Research communication)
Smooth muscles 140 (Review article)
Solid chloride ion transmitter bridge 871 (Research communication)
Solid-phase radioimmunoassay 48 (Research article)
Solid-state fermentation 263 (Research communication)
Soft matter 151 (Research article)
South Delhi fold belt 252 (Research article)
Space flight tracking and data network 410 (General article)
Sporangiate 400 (Scientific correspondence)
Sr isotopes 171 (Research communication)
Stars 193 (Generalia)
Sterculia urens 809 (Scientific correspondence)
Stereotypies 310 (Article, special section)
Stochastic differential equations 231 (General article)
Storing river runoff 227 (Scientific correspondence)
Streptozotocin 461 (Research communication)
Strong motion 772 (Research communication)
Subjective element 783 (Generalia)
Suicide behaviour 491 (Research news)
Sulphate aerosols 752 (Research article)
Summer monsoon 752 (Research article)
Superstrings 499 (General article)
Surface and interface 526 (Research communication)
Surinamite 763 (Research communication)
Symmetry methods 566 (News)
Synchrotron 572 (General article)

Taxus baccata 971 (Scientific correspondence)
Tectonic evolution 610 (Research communication)
Tectonic uplift 1023 (Research communication)
Temperature and precipitation extremes 39 (General article)
Testosterone biosynthesis 545 (Research communication)
Tilapia male 545 (Research communication)
Tomato 94 (Scientific correspondence)
Total solar eclipse 486 (News)
Traditional ecological knowledge 569 (Scientific correspondence)
Traditional Indian astronomy 705 (Historical commentary and notes)
Transmission probability 231 (General article)
Tropical diseases 946 (Book review)

Ultrafast chemical reactions 129 (Review article)

Universities, research 221 (Opinion)
University research 633 (Correspondence)
University system 802 (Commentary)
Urban development 454 (Research communication)
Use of animals in research 301 (Article, special section)
Uttarkashi earthquake 161 (Research article); 781 (Book review)

Vaccination 900 (Scientific correspondence)
Value judgement 783 (Generalia)
Vegetarianism 415 (General article)
Vertebraria 228 (Scientific correspondence)
Vitamin C 715 (In this issue); 747 (Review article); 957 (Correspondence)

Wave particle duality 426 (Review article)
Wood-chewing, horse 310 (Article, special section)
Wheat 621 (Research communication)
Wigner, Eugene Paul 375 (Historical commentary and notes)
Wild dogs 903 (Research news)
Wound healing 809 (Scientific correspondence)
Wuchereria bancrofti 257 (Research communication)

X-ray reflectometry 526 (Research communication)

Young scientists 893 (Correspondence)
Yttrium 270 (Research communication)

Zeolites 151 (Research article)
Zircon 941, 1008 (Research communications)
Zirconium 270 (Research communication)

CURRENT SCIENCE

A fortnightly journal of research

Editors: S. Ramaseshan and P. Balaram

Editorial Board

S. Arunachalam, Central Electrochemical Research Institute, Karaikudi 623 006
Asis Datta, School of Life Sciences, Jawaharlal Nehru University, New Delhi 110 067
D. Balasubramanian, Centre for Cellular and Molecular Biology, Hyderabad 500 007
J. Chandrasekhar, Department of Organic Chemistry, Indian Institute of Science, Bangalore 560 012
R. Gadagkar, Centre for Ecological Sciences, Indian Institute of Science, Bangalore 560 012
K. N. Ganeshaiah, Department of Plant Genetics and Breeding, University of Agricultural Sciences, Bangalore 560 065
R. V. Hosur, Tata Institute of Fundamental Research, Bombay 400 005
N. V. Joshi, Centre for Ecological Sciences, Indian Institute of Science, Bangalore 560 012
Jyotirmoy Das, Indian Institute of Chemical Biology, Calcutta 700 032
C. C. Kartha, Division of Cellular and Molecular Cardiology, Sree Chitra Tirunal Institute for Medical Sciences and Technology, Thiruvananthapuram 695 011
S. S. Krishnamurthy, Inorganic and Physical Chemistry Department, Indian Institute of Science, Bangalore 560 012
S. C. Lakhotia, Department of Zoology, Banaras Hindu University, Varanasi 221 005
K. Muralidhar, Department of Zoology, Delhi University, Delhi 110 007
R. Narasimhan, CMC Limited, Bangalore 560 001
K. Neelakantan, ANURAG (Advanced Numerical Research and Analysis Group), Hyderabad 500 258
R. Nityananda, Raman Research Institute, Bangalore 560 080
T. Padmanabhan, Inter-University Centre for Astronomy and Astrophysics, Pune 411 007
B. R. Pai, Propulsion Division, National Aerospace Laboratories, Bangalore 560 017
T. J. Pandian, School of Biological Sciences, Madurai Kamaraj University, Madurai 625 021
Y. S. Rajan, Department of Science and Technology, New Delhi 110 016
M. Ramakrishnan, Geological Survey of India, Calcutta 700 016
A. V. Ramani, T. T. K. Pharma Ltd., Bangalore 560 025
S. Ramaswamy, Department of Physics, Indian Institute of Science, Bangalore 560 012
A. K. Raychaudhuri, Department of Physics, Indian Institute of Science, Bangalore 560 012
S. R. Shetye, Physical Oceanography Division, National Institute of Oceanography, Dona Paula, Goa 403 004
V. Siddhartha, 51 Bharati Nagar, New Delhi 110 003
V. Srinivas, School of Mathematics, Tata Institute of Fundamental Research, Bombay 400 005
G. Srinivasan, Raman Research Institute, Bangalore 560 080
R. Uma Shaanker, Department of Crop Physiology, University of Agricultural Sciences, Bangalore 560 065
K. S. Valdiya, Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore 560 064
G. Venkataraman, Department of Physics, Sri Sathya Sai Institute of Higher Learning, Prasanthi Nilayam 515 134
M. Vidyasagar, Centre for Artificial Intelligence and Robotics, Bangalore 560 001
K. VijayRaghavan, TIFR Centre, Indian Institute of Science Campus, Bangalore 560 012

Current Science, founded in 1932, is published by the Current Science Association in collaboration with the Indian Academy of Sciences. The journal is also intended as a medium for communication and discussion of important issues that concern science and scientific activity. All articles published in *Current Science*, especially editorials, opinions and commentaries, letters and book reviews, are deemed to reflect the individual views of the authors and not the official points of view, either of the Current Science Association or of the Indian Academy of Sciences.

Working Committee, Current Science Association

P. Balaram, Indian Institute of Science, Bangalore
R. Chidambaram, Department of Atomic Energy, Bombay
V. K. Gaur, CSIR Centre for Mathematical Modelling and Computer Simulation, Bangalore
K. Kasturirangan, ISRO Satellite Centre, Bangalore
P. K. Kaw, Institute for Plasma Research, Gandhinagar
N. Kumar, Raman Research Institute, Bangalore (*Vice-President*)
N. V. Madhusudana, Raman Research Institute, Bangalore (*Secretary*)
A. Mani, Field Research Unit, Indian Institute of Tropical Meteorology, Bangalore (*Past-President*)
R. A. Mashelkar, Council of Scientific and Industrial Research, New Delhi
R. Narasimha, Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore
G. Padmanaban, Indian Institute of Science, Bangalore (*Vice-President*)
R. Ramachandran, The Institute of Mathematical Sciences, Madras
P. Rama Rao, Defence Research and Development Organization, New Delhi
S. Ramaseshan, Raman Research Institute, Bangalore (*Treasurer*)
M. S. Valiathan, Manipal Academy of Higher Education, Manipal (*President*)

Editorial Office

Editorial staff

Chandrika Ramesh, G. Madhavan, T. D. Mahabaleswara, C. S. Ravi Kumar, M. S. Venugopal

Circulation and Accounts

Peter Jayaraj, Ranjini Mohan, B. Sethumani, Shanthi Bhaskar, B. K. Shivaramiah

Annual subscription 1995

	Institutional	Personal
India	Rs 350	Rs 100
All other countries (AIR MAIL)	US $175	US $50

Single copies other than special issues: Rs 35/US $15

Editorial Office: CURRENT SCIENCE
C. V. Raman Avenue
P.B. 8001
Bangalore 560 080, India
Telephone: 91-80-3342310
Telex 845-2178 ACAD IN
Fax: 91-80-3346094

Copyright © 1995, Indian Academy of Sciences

Regd. No. KRN A15

ISSN: 0011-3891

Resonance

journal of science education

Chief Editor
N Mukunda

Resonance is a new monthly journal of science education, designed to enrich learning and teaching of science, particularly at the undergraduate level. The material presented should prove valuable in the assimilation and communication of important concepts and facts, and will include suggestions for experiments and observation. General articles, guest columns, series devoted to specific themes, research news, book reviews, reprinted classics, questions and answers, a teachers' section, poster centre-folds, quotations and anecdotes, and career information will all be included. Each issue will be well-illustrated and have a balanced coverage of different areas of science.

Sample Contents

Origin (?) of the Universe	*J V Narlikar*
The Earth's Changing Climate	*P K Das*
Lasers	*K C Rustagi*
Fermat's Last Theorem	*C S Yogananda*
Geometry	*K Paranjape*
Randomness and Probability	*R L Karandikar*
Fascinating Organic Transformations	*S Ranganathan*
Resonance in Chemistry	*J Chandrasekhar*
Life: Complexity and Diversity	*Madhav Gadgil*
Know your Chromosomes	*Vani Brahmachari*
Honeybee Dance Language	*R Gadagkar*
Teaching Quantum Mechanics	*B M Udgaonkar*
Learning Organic Chemistry through Natural Products	*N R Krishnaswamy*

- *Features:* ● *Molecule of the Month* ● *Nature Watch* ● *What is New in Computers*
- *Research News:* ● *Top Quark Discovery.* ● *Bose-Einstein Condensation*



Published by
Indian Academy of Sciences
Starting January 1996

Annual Subscription Rates

	Personal	Institutional
India	Rs 100	Rs 200
Third World Countries	US $ 25	US $ 50
Other Countries	US $ 50	US $ 100

Send your subscriptions by cheque, DD or MO in favour of "Indian Academy of Sciences"

Indian Academy of Sciences, C V Raman Avenue, PB 8005, Bangalore 560 080, India. Tel: +91 (80) 334 2546, Fax: +91 (80) 334 6094